

STIC-Biotech/Ch mLib

73941

From: Rao, Manjunath N.
Sent: Thursday, August 22, 2002 4:49 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/870113

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

Date: 8-22-02

Please search the following as soon as possible for application with serial number **09/870,113**

SEQ ID NO: 1 against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

SEQ ID NO: 2, 4, 8, 10, 12 against all commercial protein databases including issued patents database and pending application database and provide a **print of all results**.

If y u have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biot chnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

RECEIVED
AUG 22 2002
STIC

Searcher: D. Schreiber
Phone: 308-4292
Location: CM1 6A03
Date Picked Up: 8/26
Date Completed: 8/29
Searcher Prep/Review: 10
Clerical: _____
Online time: 7

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: CompuGen
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

35

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:57:53 ; Search time 130.92 Seconds
(without alignments)
480.982 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGGAGGAPG.....VVEFFKYLTKRQEWRAKG 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_protein:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1940	100.0	364	Q96A46	Q96A46 homo sapien
2	1176	60.6	338	11 Q920G8	Q920G8 mus musculus
3	1013	52.2	347	4 Q9NYZ2	Q9NYZ2 homo sapien
4	937	48.3	177	4 Q969G8	Q969G8 homo sapien
5	910	46.9	176	4 Q9H2J3	Q9H2J3 homo sapien
6	799	41.2	312	5 Q23125	Q23125 caenorhabdi
7	781	40.3	379	5 Q9VAY3	Q9VAY3 drosophila
8	742.5	38.3	303	5 Q94638	Q94638 onchocerca
9	725.5	37.4	187	4 Q9P0J2	Q9P0J2 homo sapien
10	722.5	37.2	187	11 Q9CQG7	Q9CQG7 mus musculus
11	717.5	37.0	380	5 Q9NHV6	Q9NHV6 drosophila
12	672.5	34.7	301	5 Q94634	Q94634 onchocerca
13	601	31.0	303	3 Q14281	Q14281 schistosach
14	521	26.9	781	10 Q9LMD6	Q9LMD6 arabidopsis
15	510	26.3	289	10 Q82049	Q82049 ribes nigru
16	507	26.1	331	10 Q64731	Q64731 arabidopsis

17	444.5	22.9	155	4	Q969S1	Q969S1 homo sapien
18	444.5	22.9	182	11	Q912Y0	Q912Y0 mus musculus
19	308.5	15.9	413	5	Q9VWF9	Q9VWF9 drosophila
20	305.5	15.7	449	5	Q9V105	Q9V105 drosophila
21	303	15.6	320	11	Q921P8	Q921P8 mus musculus
22	300	15.5	325	10	Q94AG6	Q94AG6 arabidopsis
23	297	15.3	367	10	Q9C910	Q9C910 arabidopsis
24	296	15.3	321	4	Q9BSK2	Q9BSK2 homo sapien
25	294.5	15.2	358	5	Q9VQ37	Q9VQ37 drosophila
26	293	15.1	351	4	Q9Y379	Q9Y379 homo sapien
27	293	15.1	351	4	Q9UF66	Q9UF66 homo sapien
28	291	15.0	365	5	Q9STJ5	Q9STJ5 drosophila
29	289	14.9	359	4	Q9BZJ4	Q9BZJ4 homo sapien
30	284.5	14.7	706	3	Q9C2K1	Q9C2K1 neurospora
31	282.5	14.6	902	3	Q12482	Q12482 saccharomyc
32	282	14.5	412	10	Q9FHX2	Q9FHX2 arabidopsis
33	274	14.1	269	5	Q18934	Q18934 caenorhabdi
34	274	14.1	330	10	Q9SVB2	Q9SVB2 arabidopsis
35	272	14.0	337	4	Q9P182	Q9P182 homo sapien
36	271	14.0	297	5	Q9VBN7	Q9VBN7 drosophila
37	270	13.9	338	4	Q9UHR1	Q9UHR1 homo sapien
38	269	13.9	676	11	Q9DCF5	Q9DCF5 mus musculus
39	268	13.8	409	11	Q91YK1	Q91YK1 mus musculus
40	268	13.8	675	11	Q9CZF6	Q9CZF6 mus musculus
41	267.5	13.8	271	3	Q60029	Q60029 ashbya goss
42	266	13.7	311	4	Q96CQ1	Q96CQ1 homo sapien
43	266	13.7	359	11	Q9D8K8	Q9D8K8 mus musculus
44	265.5	13.7	310	4	Q9NVN5	Q9NVN5 homo sapien
45	262	13.5	307	4	Q14589	Q14589 homo sapien

ALIGNMENTS

RESULT 1

Q96A46	PRELIMINARY;	PRT;	364 AA.
ID Q96A46			
AC Q96A46;			
DT 01-DEC-2001 (TrEMBLrel. 19, Created)			
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE			
DE CARRIER SPLICING VARIANT).			
GN HMRS3/4.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP MEDLINE=21195335; PubMed=11297739;			
RX Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,			
RA Larsson C., Suomalainen A.;			
RT "Characterization of a novel human putative mitochondrial transporter			
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";			
RL FEBS Lett. 494:79-84(2001).			
DR EMBL; AJ303077; CAC27996.1; -			
DR EMBL; AF327402; AAK49519.1; -			
SQ SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;			

Query Match 100.0%; Score 1940; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.3e-159; Indels 0; Gaps 0;
Matches 364; Conservative 0; Mismatches 0;

QY 1 MELEGRGAGGAGGAPGSPGESALLDGLQGRVGRGAGGAGACRPPVRRQPDG 60

|||||

Db 1 MELEGRGAGGAGGAPGSPGESALLDGLQGRVGRGAGGAGACRPPVRRQPDG 60

|||||

QY 61 PDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNLEALWR 120

|||||

Db 61 PDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNLEALWR 120

Post dated

```
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVHPGNSHIANGAAGCVA 180
DB 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVHPGNSHIANGAAGCVA 180
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
DB 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
QY 241 IHFTYEFLOEHFNPRPNPSSHVLSGACAGAVAAATPLDVCCKTLTNTQESIALNSH 300
DB 241 IHFTYEFLOEHFNPRPNPSSHVLSGACAGAVAAATPLDVCCKTLTNTQESIALNSH 300
QY 301 ITGHTITMASAFRTYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEW 360
DB 301 ITGHTITMASAFRTYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEW 360
QY 361 RAGK 364
DB 361 RAGK 364

RESULT 2
Q920G8      PRELIMINARY;      PRT;      338 AA.
ID Q920G8;
AC Q920G8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RA She J.-X.;
RT "A new gene which is highly expressed in NOD mice spleen.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF288621; AAL23859.1; -.
SQ SEQUENCE 338 AA; 37510 MW; BB35B1F70C56A3FE CRC64;
```

```
Query Match 60.6%; Score 1176; DB 11; Length 338;
Best Local Similarity 61.9%; Pred. No. 1.4e-93;
Matches 224; Conservative 48; Mismatches 62; Indels 28; Gaps 4;

QY 1 MELEGRAGGAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACRPPVQDDPSG 60
DB 1 MELRRGGVGNAAAG-----RRMDGDCRDG---GCGSKDAGS----- 33
QY 61 PDYALPAGATVTTMVGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYNNVLEALWR 120
DB 34 EDYENLPTSASVSTHTAGAMAGILEHSIMYPVDSVKTSMOSLPDPKARYTSIYGALKR 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVHPGNSHIANGAAGCVA 180
DB 94 IMHTEGFWRPLRGLNVMVMGAGPAHAMYFACYENKRTLNDVFSHQNSHLANGVAGSMA 153
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
DB 154 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
QY 241 IHFTYEFLOEHFNPRPNPSSHVLSGACAGAVAAATPLDVCCKTLTNTQESIALN-S 299
DB 214 IHFTYEFLOEQVNPRTYNPQSHIISGLAGALAAATPLDVCCKTLTNTQESIALN-SLA 273
QY 300 HITGHITGMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 359
DB 274 NVSGRLSMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 333
QY 360 WR 361
```

```
Db 334 NR 335

RESULT 3
Q9NYZ2      PRELIMINARY;      PRT;      347 AA.
ID Q9NYZ2;
AC Q9NYZ2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF223466; AAF64141.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 52.2%; Score 1013; DB 4; Length 347;
Best Local Similarity 60.8%; Pred. No. 1.8e-79;
Matches 197; Conservative 41; Mismatches 58; Indels 28; Gaps 5;

QY 1 MELEGRAGGAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACRPPVQDDPSG 60
DB 1 MELR---SGSV---GSQAVARRMDGDS-----RDGGGG-----KDATGS 33
QY 61 PDYALPAGATVTTMVGAVAGILEHCVMYPIDCVKTRMOSLOPDPAARYNNVLEALWR 120
DB 34 EDYENLPTSASVSTHTAGAMAGILEHSIMYPVDSVKTSMOSLPDPKARYTSIYGALKR 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSLSDVHPGNSHIANGAAGCVA 180
DB 94 IMRTEGFWRPLRGLNVMVMGAGPAHAMYFACYENKRTLNDVFSHQNSHLANGVAGSMA 153
QY 181 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
DB 154 TLLHDAAMNPAEVVKQRMQYNSPYHRVTDVAVWQNEGAGAFYSYTTTQLTMTNVPFQA 240
QY 241 IHFTYEFLOEHFNPRPNPSSHVLSGACAGAVAAATPLDVCCKTLTNTQESIALN-S 299
DB 214 IHFTYEFLOEQVNPRTYNPQSHIISGLAGALAAATPLDVCCKTLTNTQESIALN-SLA 273
QY 300 HITGHITGMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 323
DB 274 NVSGRLSMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 297
QY 300 HITGHITGMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 323
DB 274 NVSGRLSMAFAFRVYQVGGVYATFRGQARVIYQIPSTAIWSVYEFKYLITKRQEE 297

RESULT 4
Q969G8      PRELIMINARY;      PRT;      177 AA.
ID Q969G8;
AC Q969G8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE
DE CARRIER SPLICING VARIANT).
GN HMRS3/4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```


RP SEQUENCE FROM N.A.
RX MEDLINE=21195335; PubMed=11297739;
RA Li F., Nikali K., Grehan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";
RL FEBS Lett. 494:79-84(2001).
DR EMBL: AJ303078; CAC27997.1; -
DR EMBL: AF327403; AAK49520.1; -
SQ SEQUENCE 177 AA; 20045 MW; E5370283AAF732E3 CRC64;

Query Match 48.3%; Score 937; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.7e-73;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 MNPAEVVKORMQYNPSYHRVTDVAVWQNEGAGAFYSYTTQLTMNVPFOAIHFMTYE 247
Db 1 MNPAEVVKORMQYNPSYHRVTDVAVWQNEGAGAFYSYTTQLTMNVPFOAIHFMTYE 60
QY 248 FLOEHFNPPRRYNSHVLSCACAGAVAAATPLDVCKTLNTQESLALNSHITGHTG 307
Db 61 FLOEHFNPPRRYNSHVLSCACAGAVAAATPLDVCKTLNTQESLALNSHITGHTG 120
QY 308 MASAFRTYQVGGVTAIFRGVQARVIYQIPSTAIWSVYEFFKYLITKROEWRAGK 364
Db 121 MASAFRTYQVGGVTAIFRGVQARVIYQIPSTAIWSVYEFFKYLITKROEWRAGK 177

RESULT 5
Q9H2J3 PRELIMINARY; PRT; 176 AA.
ID Q9H2J3
AC Q9H2J3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NP0016.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PITUITARY;
RA Xu X., Yang Y., Gao G., Xiao H., Chen Z., Han Z.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF267854; AAG44723.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 2.
SQ SEQUENCE 176 AA; 19931 MW; 4A4831AC572309D4 CRC64;

Query Match 46.9%; Score 910; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 5.9e-71;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VVKORMQYNPSYHRVTDVAVWQNEGAGAFYSYTTQLTMNVPFOAIHFMTYEFLOEH 252
Db 5 VVKORMQYNPSYHRVTDVAVWQNEGAGAFYSYTTQLTMNVPFOAIHFMTYEFLOEH 64
QY 253 FNPPRRYNSHVLSCACAGAVAAATPLDVCKTLNTQESLALNSHITGHTGMSAF 312
Db 65 FNPPRRYNSHVLSCACAGAVAAATPLDVCKTLNTQESLALNSHITGHTGMSAF 124
QY 313 RTYQVGGVTAIFRGVQARVIYQIPSTAIWSVYEFFKYLITKROEWRAGK 364
Db 125 RTYQVGGVTAIFRGVQARVIYQIPSTAIWSVYEFFKYLITKROEWRAGK 176

RESULT 6
Q23125 PRELIMINARY; PRT; 312 AA.
ID Q23125
AC Q23125;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z66521; CAA91399.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 41.2%; Score 799; DB 5; Length 312;
Best Local Similarity 49.5%; Pred. No. 5e-61;
Matches 159; Conservative 54; Mismatches 82; Indels 26; Gaps 5;

QY 42 GGGEAGACGPPVRQDPDGPDEALPAGATVTHWAGAVAGILCHCVMPIDCVKTRMQ 101
Db 3 GGGE-----DEYESLPT-HSVPHLTAGALAGAVEHCVMFPFDSVKTRMQ 46
QY 102 SLOPDPAARYRNVLALRIIRTEGLWRMRGLNVTATGAGPAHALYFACYEKLKLTSD 161
Db 47 SLPCPETKCTPVHSLMSIVKREGRLPLRGVNAAGSMPAHALYFYVEKMKGLTG 106
QY 162 VIHPGGNSH-IANGACCVATLLHDAAMNPAEVVVKORMQYNPSYHRVTCVRAVWNEG 220
Db 107 --NSAGHSNTLAYGAGVATLIHDAIMNPAEVVVKORMQAFSPYGSLSLECARCVNREG 164
QY 221 AGAFYSYTTQLTMNVPFOAIHFMTYEFLOEHFNPPRRYNSHVLSCACAGAVAAAT 280
Db 165 VAAYFSYTTQLAMNVPFOAIHFMSYEFWOHVLNPEHKYDPKSHLIAGGLAGLAAALT 224
QY 281 PLDVCKTLNTQESLALNS-----HITGHITGMSAFRTYQVGGVTAIFRGVQARVI 333
Db 225 PMDCVKTVLNTQQAADPANRRIFLQARYRYRGISDAVRTIYSQRLSGFSGCGLQARVI 284
QY 334 YQIPSTAIWSVYEFFKYLIT 354
Db 285 FQVPATALSWSVYELFKFMLS 305

RESULT 7
Q9VAY3 PRELIMINARY; PRT; 379 AA.
ID Q9VAY3
AC Q9VAY3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG4963 PROTEIN (GH09840P).
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Suton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Ketchum K.A.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Wellsteinbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin J., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC -!- INNER MEMBRANE (BY SIMILARITY).
 CC EMBL; AE003763; AAF56764.1; -
 DR EMBL; AY060268; AAL25307.1; -
 DR FlyBase; FBgn0039561; CG4963.
 DR InterPro; IPR001993; Mitochondrion.
 DR InterPro; IPR002067; Mitochondrion.
 DR Pfam; PF00153; mito_carrier.
 DR PRINTS; PR00926; MITOCH_CARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 2.
 KW Inner membrane; Mitochondrion; Transmembrane; Transport.
 SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 40.3%; Score 781; DB 5; Length 379;
 Best Local Similarity 52.7%; Pred. No. 2.3e-59;
 Matches 154; Conservative 41; Mismatches 83; Indels 14; Gaps 4;

Qy 62 DYALPAGATVTHMVAGAVAGILEHCVMYPTDCVKTQMSLQPDPAARVNLVLEALWRI 121
 Db 5 DYESLPT-TSVCVNTAGIAGVLEHVVMYPLDSVKTQMSL--SPPTKNNIYVTLRTM 61
 Qy 122 IRTGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLDSDVHPGNSHIANGAGCVAT 181
 Db 1 IRTGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLDSDVHPGNSHIANGAGCVAT 181

Db 62 ITRGLLRPIRGASAVVLGAGPAHSLYFAAYEMTKELAKFTSVRLNLYVISGA---VAT 118
 Qy 182 LLHDAAMNPAEVVKORMOMYNSPHYRVTDVAVRVQWNGAGAFYRSYTOTLTNNVFOAI 241
 Db 119 LIHDAISSPTDVIKORMOMYNSPYTSVVCVRDIYKRGFKAFYRAYGTQLVWNLPIYOTI 178
 Qy 242 HFMTYEFQEHFNPPORRYNSPSHVLGACAGAAAAATTPLDVCKTLLTNTQESLALNSHI 301
 Db 179 HFITYEFFQNMNLKERNPPVHMAAGAAAGACAAVTPDLVDIKTLNTOE----- 230
 Qy 302 TGHITGASAFVTVQVGVATYAFRGVQARVYIQIPSTAIAMSVYEFKYL 353
 Db 231 TGLTRGMEASRKIYHMAAGLGLFRGTTRVLYSMPTAICWSTYEFKYL 282
 RESULT 8
 Q94638 ID PRELIMINARY; PRT; 303 AA.
 AC Q94638;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MITOCHONDRIAL SOLUTE CARRIER.
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96326580; PubMed=8703971;
 RA Catull J., Miller D.J.;
 RT "cDNAs from *Onchocerca* sp. encoding members of the MRS3/MRS4 class of
 RT Mitochondrial solute carriers.";
 RL Biochim. Biophys. Acta 1282:179-181(1996).
 DR EMBL; U45998; AAB19037.1; -
 DR InterPro; IPR001993; Mitochondrion.
 DR Pfam; PF00153; mito_carrier.
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
 SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 38.3%; Score 742.5; DB 5; Length 303;
 Best Local Similarity 45.2%; Pred. No. 3.7e-56;
 Matches 142; Conservative 54; Mismatches 87; Indels 31; Gaps 3;

Qy 49 CRPPVRQDPDGGPDYALPAGATVTHMVAGAVAGILEHCVMYPTDCVKTQMSLQPDPA 108
 Db 12 CRWPV-----HLTLAGSVAGLAHCLMFPDPSVKTRLSQSLCPCE 50
 Qy 109 ARYRNLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLDSDVHPGNS 168
 Db 51 TSCPTAMHSLMSVMKREGLLSLKGVNAVLTIPAHAFYTVYENSAYLLN--NPRVS 108
 Qy 169 SHIANGAAGCVATLLHDAAMNPAEVVKORMOMYNSPHYRVTDVAVRVQWNGAGAFYRSY 228
 Db 109 NSVSAISGALATVTHDAVMNPAEVVKORMOMIFSPKQMSLEICIRYIREGLAFYRSY 168
 Qy 229 TQLTNNVFPQAIHMTYEFQEHFNPPORRYNSPSHVLGACAGAAAAATTPLDVCKTL 288
 Db 169 ITQTLNVPYOCYTHPMIYEQNLNPHHDYNPSSHLSVSGGTAGGIAAATTPDLCVKTV 228
 Qy 289 LNTQES-----LALNSHITGHITGMASAFRTVYQGVGVATYAFRGVQARVYIQIPSTA 340
 Db 229 LNTQTPRENTYRLTTOSEHTAYKGLADGKIYIYLRGTGGFGRGLQARIIPSTA 288
 Qy 341 IAWSYEFKYLIT 354
 Db 289 LWSAYELCKYMLS 302
 RESULT 9
 Q9P0J2 ID PRELIMINARY; PRT; 187 AA.

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:34:08 ; Search time 138.54 Seconds
(without alignments)
291.835 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGVAGGPAAGPG.....VYEFKYLITKROEWRAK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1696	87.4	318	22	AA41505 Human polypeptide
2	1534	79.1	289	21	AA50383 Human uncoupling p
3	1534	79.1	289	22	AA60113 Human transport pr
4	1387	71.5	268	22	AA39719 Human polypeptide
5	1113.5	57.4	366	22	AA000938 Human bone marrow
6	1025.5	52.9	272	21	AA42980 Human ORFX ORF2744
7	1012.5	52.2	331	22	AA60658 Human mitochondria
8	919	47.4	677	22	AAU29748 Novel human secret
9	916	47.2	174	22	AAU01989 Human secreted pro
10	910	46.9	176	21	AA40634 Human ORFX ORF398
11	910	46.9	176	22	AA88221 Human mitochondria

12	856.5	44.1	299	22	AA79039 Human protein SEQ
13	781	40.3	379	22	AB61130 Drosophila melanog
14	739.5	38.1	199	22	AA80023 Human protein SEQ
15	725.5	37.4	187	22	AA93909 Human polypeptide,
16	725.5	37.4	187	22	AA98222 Human mitochondria
17	725.5	37.4	187	22	AA94965 Human polypeptide,
18	717.5	37.0	187	22	AA93934 Human secreted pro
19	717.5	37.0	188	21	AA32112 O. volutus mitoch
20	558	28.8	215	22	AA49666 Human secreted pro
21	541	27.9	194	21	AA32111 Gene #25 human sec
22	515.5	26.6	181	22	AAU01988 Blackcurrant Rib7
23	510	26.3	289	18	AA17054 Arabidopsis thalia
24	509	26.2	331	21	AA22077 Arabidopsis thalia
25	507	26.1	331	21	AA43092 Arabidopsis thalia
26	497	25.6	290	21	AA22078 Arabidopsis thalia
27	495	25.5	290	21	AA43093 Arabidopsis thalia
28	453	23.4	278	21	AA22079 Arabidopsis thalia
29	451	23.2	278	21	AA43094 Arabidopsis thalia
30	440.5	22.7	155	21	AA42966 Human ORFX ORF2730
31	412	21.2	111	21	AA42968 Human ORFX ORF2732
32	343.5	17.7	149	21	AA42964 Human ORFX ORF2728
33	309	15.9	74	22	ABG01167 Novel human diase
34	308.5	15.9	413	22	AB69108 Drosophila melanog
35	305.5	15.7	449	22	AB59263 Drosophila melanog
36	300	15.5	301	21	AA52781 Arabidopsis thalia
37	300	15.5	301	21	AA52783 Arabidopsis thalia
38	300	15.5	325	21	AA52780 Arabidopsis thalia
39	300	15.5	325	21	AA52782 Arabidopsis thalia
40	300	15.5	361	21	AA52779 Arabidopsis thalia
41	294.5	15.2	358	22	AB71142 Drosophila melanog
42	293	15.1	351	21	AA50378 Human uncoupling p
43	281.5	14.5	374	22	AA40959 Human polypeptide
44	281.5	14.5	374	22	AA40960 Human polypeptide
45	281	14.5	365	21	AA42783 Human ORFX ORF2547

ALIGNMENTS

RESULT 1

AA41505
ID AA41505 standard; Protein; 318 AA.

XX AA41505;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6436.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0489725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAI60661.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6436; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 318 AA;
SQ

Query Match 87.4%; Score 1696; DB 22; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.6e-163;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GACPPVRQPDGPDYALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLOPD 106
DB 1 gacppvrqpdsgpdyaelpagatvcthmavagavagilehcvmpidcvktrmqslqpd 60

QY 107 PAARYNVLEALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTKTSLSDVIHPG 166
DB 61 paarynvlealwriirteglrpmrglnvtatgagpahalyfacyeklkktlsdvihpg 120

QY 167 GNSHIANGAGCVATLLHDAAMPVAVVKKORMQYNPSYHRVTDVAVVQNEGAGAFYR 226
DB 121 gnschiangagcvatlilhdaampvavvkgqrmqynpsyhrvtdcvravvqnegagafyr 180

QY 227 SYTTQTMNVVFOAIIHMTYEEFLQEHFNPPORRYNPSSHVLSCACAGAVAAATPLDVCK 286
DB 181 syttqlcmnvvpfoaiahmtyeeflqehfnpporrynpsshvlsacagavaaaatpldvck 240

QY 287 TLLNTQESLALNSHITGHITGMSAFRTVYQVGVTAYFRGQVARYIQTIPSTAIWSVY 346
DB 241 tllntqeslalnshithitgmsaftrvyqvgvtayfrgqvaryiqtipstaiawsy 300

QY 347 EPPKYLITKQEWRAK 364
DB 301 effkylitkrqeewragk 318

RESULT 2
AAB50383
ID AAB50383 standard; Protein; 289 AA.
XX
XX AAB50383;
XX
XX 12-MAR-2001 (first entry)
XX
XX Human uncoupling protein #6.
XX

Human; uncoupling protein; immunosuppressive; antiarthritis;
antirheumatic; antiproliferative; cardiant; vasotropic;
cerebroprotective; neuroprotective, antibacterial; ophthalmological;
gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;
gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
infertility.
XX
XX Homo sapiens.
XX OS
XX WO2000061614-A2.
XX PN
XX 19-OCT-2000.
XX PD
XX
XX 06-APR-2000; 2000WO-US09534.
XX PF
XX 09-APR-1999; 99US-0128701.
XX PR
XX 08-JUL-1999; 99US-0142821.
XX PR
XX 18-AUG-1999; 99US-0149448.
XX PR
XX 12-NOV-1999; 99US-0164751.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
XX PI
XX WPI: 2000-656322/63.
XX DR
XX N-PSDB; AAC90457.
XX
XX Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
XX Claim 11; Page 323-324; 343pp; English.
XX
XX The present sequence is a human uncoupling protein. The nucleotide
CC sequences encoding the uncoupling proteins may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
XX Sequence 289 AA;
SQ

Query Match 79.1%; Score 1534; DB 21; Length 289;
Best Local Similarity 100.0%; Pred. No. 3.8e-147;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLOPDPAARYNVLEALWRIIRTEGLRPMRGLN 135
DB 1 mvagavagilehcvmpidcvktrmqslqpdpaarynvlealwriirteglrpmrgln 60

QY 136 VTATGAGPAHALYFACYEKLKTKTSLSDVIHPGNSHANGAGCVATLLHDAAMPVAVV 195
DB 61 vtatgagpahalyfacyeklkktlsdvihpgnshiangagcvatlilhdaampavvk 120

QY 196 QRMOMYNSPYHRVTDVAVVQNEGAGAFYRSYTTQTMNVVFOAIIHMTYEEFLQEHFNP 255
DB 121 qrmqynpsyhrvtdcvravvqnegagafyrsyttqtmnvvpfoaiahmtyeeflqehfnp 180

QY 256 QRRYNPSSHVLSCACAGAVAAATPLDVCKTLLNTQESLALNSHITGHITGMSAFRTV 315
DB 181 qrrynpsshvlscacagavaaaatpldvcktllnqeslalnshithitgmsaftrv 240

QY 316 YQVGGVTAYFRGQVARYIQTIPSTAIWSVYEFFKYLITKQEWRAK 364
DB 241 yqvgvtayfrgqvaryiqtipstaiawsyeffkylitkrqeewragk 289

RESULT 3
AAB60113

ID AAB60113 standard; Protein; 289 AA.
AC AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2000078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI; 2001-041424/05.
DR N-PSDB; AAF27733.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 2; Page 133-134; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated tppts). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
XX Sequence 289 AA;

Query Match 79.1%; Score 1534; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. No. 3.8e-147;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 MVAGAVAGILEHCVMYPIDCVKTRMOSLOPDPAPARYNVLEALWRIIRTEGLWRPVRGLN 135
Db 1 mvagavagilehcvmypidcvktrmqlgdpdparynrvlealwriirteglwrpvrgrln 60
QY 136 VTATGAGPAHALYFACYEKIKTKLSDVIHPGGNSHTANGACGCVATLLHDAAMNPAEVVK 195
Db 61 vtatgagpahalyfacyekiktklsdvihpggnshtangagcgvatllhdaamnpaevvk 120
QY 196 QRMQNTNSPHYRVTDCVRVWQNEGAGAFYSYTTQLTMMVPFOAHFMTYEFLOQHFN 255
Db 121 qrmqntnsphyrvtdcavrwnqnegagafysyttltmmvpfoahfmyeflqhfnfp 180
QY 256 QRRYNPSSHVLGACAGAVAAATTPDLVCKTLNLTQESLALNSHITGHITGMASAFRTV 315
Db 181 qrrynpsshvlsgacagavaaattpldvcktlntlntqeslalnshitghitgmasafrtv 240
QY 316 YQVGGVTAYFRGVQARVYIQIPSTAIWSVYEFFKYLITRQEWRAK 364
Db 241 yqvggvtayfrgvqarvyiqipstaiwsvyeffkylitrqewragk 289

RESULT 4
AAM39719
ID AAM39719 standard; Protein; 268 AA.
XX
AC AAM39719;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2864.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI58875.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2864; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAAM2213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 268 AA;

Query Match 71.5%; Score 1387; DB 22; Length 268;
Best Local Similarity 98.1%; Pred. No. 2.9e-132;
Matches 262; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 TRMQSLQPDPAARYNRVLEALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKK 157

KW Immune suppression; Immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
OS WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US08656.
XX 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
XX Claim 20; Page 189; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 677 AA;

Query Match 47.4%; Score 919; DB 22; Length 677;
Best Local Similarity 59.2%; Pred. No. 3.4e-84;
Matches 181; Conservative 37; Mismatches 60; Indels 28; Gaps 5;
QY 1 MELEGAGGAGGAAAGPGRSPGESALLDGLQGVGRGAGGACRPPVVRQDPDSG 60
DB 13 melr---sgsv--gsqavarimdgds-----rdgggg-----kdatgs 45
QY 61 PDYCALPAGATVTHVMAGAVAGILEHCVMYPIDCVTRMQSLQDPDPAARYNVLALWR 120
DB 46 edyenltsasvsthtagmagilehsvmypdvsvtrmgslspsgspvsiygalkk 105
QY 121 IIRTEGLWRMRGILNVATGAGPAHALYFACYEKLKLTSLDVHPGNGSHNTANGACGVA 180
DB 106 imrtegfwrpgrgvnmimgagpahanmyfcyemknrtlnvdfhngnshlangagsma 165
QY 181 TLLHDAAMNPAEVVKORMQMYNSPYHRVTCVRAVWQNEGAGAFYRSTYTTQLTMNVPFOA 240
DB 166 tllhdavmnpaevvkvqlmqnysqhsraiscirtvrtgclgafyrsyttqltmnifqs 225
QY 241 IHFMTYEFLEHFNQRRYNSPSSHLVSGACAGAVAAATTPLDVCKTLLNTQESLAIN-S 299
DB 226 ihfityeflgeqvnphtynphshlisgglagalaataatpldvcktlntqenvalsla 285
QY 300 HITGHI 305
DB 286 nlsgrl 291

RESULT 9
AAU01989
ID AAU01989 standard; Protein; 174 AA.
XX AAU01989;
XX 29-AUG-2001 (first entry)
XX Human secreted protein encoded by gene #25.
DE Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX Homo sapiens.
XX WO200123598-A1.
XX 05-APR-2001.
XX 26-SEP-2000; 2000WO-US26324.
XX 27-SEP-1999; 99US-0155807.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Komatsoulis G, Ruben SM, Rosen CA;
XX WPI; 2001-281684/29.
XX N-PSDB; AAS03906.
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX Disclosure; Page 501; 518pp; English.
XX The sequence represents a human secreted protein encoded by a nucleic
CC acid of the invention. Secreted proteins and their related nucleic acids
CC can be used in the diagnosis of or susceptibility to a pathological
CC condition by determining the presence or absence of a mutation in a
CC nucleic acid or the presence or amount of expression of a secreted
CC protein. The sequences are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, used in
CC chickens or sheep. The antibodies to the polypeptides can also be used in
CC alleviating symptoms associated with disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent
CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The peptides can also
CC be used to aid wound healing and epithelial cell proliferation, to help
CC prevent skin ageing due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues, in chemotaxis and as a food
CC additive or preservative to alter storage capabilities.
XX
SQ Sequence 174 AA;

Query Match 47.2%; Score 916; DB 22; Length 174;
Best Local Similarity 100.0%; Pred. No. 1e-84;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 TLLHDAAMNPAEVVKORMQMYNSPYHRVTCVRAVWQNEGAGAFYRSTYTTQLTMNVPFOA 240
DB 1 tllhdaamnpaevvkvqrmqymnspyrhrvtcdvraavwqnegafyrsttqltmnvpfqa 60
QY 241 IHFMTYEFLEHFNQRRYNSPSSHLVSGACAGAVAAATTPLDVCKTLLNTQESLAINSH 300

|||||
Db 61 ifhmfyeqlqehfnpqrrypnsshvlgacagavaaaattpldvckllntqeslalnsh 120
QY 301 ITGHTGMSAFTYVQGVGTAYFRGVQARVIYQIPSTAIWSVYEFKYLIT 354
Db 121 itghtgmasafrtyvqgvgtayfrgvqarvlyqipstaiaawsyeffkylit 174
RESULT 10
AAB40634
ID AAB40634 standard; Protein; 176 AA.
XX
AC AAB40634;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF398 polypeptide sequence SEQ ID NO:796.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC74843.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 854; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,

CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
SQ Sequence 176 AA;
Query Match 46.9%; Score 910; DB 21; Length 176;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 VVKRQMOMNSPYHRVTDVRAVQWQEGAGAFYRSYTTQLTMNVFPQAIHFMTYELQEH 252
Db 5 vvkqrmqmnspyrhvtcdvrvavqnegagafyrsyttqltmnvfpqaihfmtyeflqeh 64
QY 253 FNPORRYNPSSHVLSGACAGAVAAAATPLDVCKLLNTQESLALNSHTTGHTGMSAF 312
Db 65 fnpqrrypnsshvlgacagavaaaattpldvckllntqeslalnshitghitgmasaf 124
QY 313 RTYVQGVGTAYFRGVQARVIYQIPSTAIWSVYEFKYLITKROEWRAGK 364
Db 125 rtyvqgvgtayfrgvqarvlyqipstaiaawsyeffkylitkrqewragk 176
RESULT 11
AAB98221
ID AAB98221 standard; Protein; 176 AA.
XX
AC AAB98221;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
XX
KW Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
XX
OS Homo sapiens.
XX
PN CN1281895-A.
XX
PD 31-JAN-2001.
XX
PF 27-JUN-2000; 2000CN-0116795.
XX
PR 27-JUN-2000; 2000CN-0116795.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Li N, Xu X, Xiao H;
XX
DR WPI; 2001-282781/30.
DR N-PSDB; AAB22162.
XX
PT New human mitochondrial solute carrier (hMSC) protein isomer and its
PT code sequence -
XX
PS Claim 4; Page 18 (disclosure); 22pp; Chinese.
XX
CC The present invention describes a human mitochondrial solute carrier
CC protein, designated hMSC-homologue, which is expressed in human
CC hypophysis. Also described are methods for the preparation and
CC detection of hMSC-homologue protein and nucleotide sequences. The
CC present sequence represents hMSC-homologue, as given in the present
XX invention.
SQ Sequence 176 AA;
Query Match 46.9%; Score 910; DB 22; Length 176;

Best Local Similarity 100.0%; Pred. NO. 4.2e-84; Mismatches 0; Indels 0; Gaps 0; Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VVKORQMNSPYHRVTDVAVQWNEGAGAFYRSYTTQLTMMVPCATHFWTYEFLQEH 252
Db 5 vvkqrmqmyspyhrvtdcvravwqnegagafyrsyttqltmmvpqahfmyeflqeh 64

QY 253 FNPQRRNPSSHVLGACAGAVAAAATTPLDVCKTLINTQESLALNSHITGHITGMASAF 312
Db 65 fnpqrrnpsshvlsgacagavaaaattpldvcktlintqeslalnshitghitgmasaf 124

QY 313 RTVYQGVGTAYPRGVOARVIYQIPSTAIAMSVEFFKYLITKRQEWRAK 364
Db 125 rtvyqgvgtayirgvgarviyqipstaiawsyeffkylitkrqeewragk 176

RESULT 12
AAM79039
ID AAM79039 standard; Protein; 299 AA.
XX AAM79039;
XX AAM79039;
DT 06-NOV-2001 (first entry)
XX Human protein SEQ ID NO 1701.
XX Human; cytokine: cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYPSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52172.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX Claim 20; Page 4043; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 299 AA;
SQ

Query Match 44.1%; Score 856.5; DB 22; Length 299;
Best Local Similarity 59.6%; Pred. No. 2.4e-78;
Matches 167; Conservative 30; Mismatches 56; Indels 27; Gaps 4;

QY 1 MELEGRGAGVAGGPAAGPCRGSPGESALLDGLQVGRGAGGAGACRPPVRQDDSG 60
Db 28 melr---sgsv--gsqavarimdgds-----rdgggg-----kdatgs 60

QY 61 PDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKRMQSLQDPDPAARYRNVLALWR 120
Db 61 edyenlptsasvsthtagamagilehsvpydvsktrmqslspdpkagytvsygaikk 120

QY 121 IIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLAKTLDSDVIHPGNSHIANGAGCVA 180
Db 121 imrtgfwrlprlgvnmimgagpahamyfacyenmkrtlnvdfhggshlangiagsma 180

QY 181 TLLHDAAMNPAEYVKORMQMYNSPYHRVTDVAVQWNEGAGAFYRSYTTQLTMMVPCFA 240
Db 181 tllhdavmnpaevvkqrlqmynsqhrsaiscirtvrtglgafyrsyttqtlmnpifqs 240

QY 241 IHEWTVEFFLQEHFNPPRRNPSSHVLGACAGAVAAAATT 280
Db 241 ihityeflqeqvnrhtynpqshlisggilagalaanaans 280

RESULT 13
ABB61130
ID ABB61130 standard; Protein; 379 AA.
XX ABB61130;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 10182.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL05233.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 10182; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 379 AA;

Query Match 40.3%; Score 781; DB 22; Length 379;
Best Local Similarity 52.7%; Pred. No. 1.5e-70;
Matches 154; Conservative 41; Mismatches 83; Indels 14; Gaps 4;

QY 62 DYEALPAGATVTHVAGAVAGILEHCVMYPIDCVKTRMQSLQPPAARYNVLALWRI 121
Db 5 dyeslpt-tsvgvnmtdaigvlehvmvypidsvktmqsl--spptkmmivstlrcm 61
QY 122 IRTEGLWRPGRGLNVTATGAGPAHALYFACYEKLKLTSDVIHPGNSHIANGAAGCVAT 181
Db 62 itregllprigasavllgagpahalyfaayemtkeltakftsvrnlmnyvisga---vat 118
QY 182 LHDAAMNPAEYVKORMQMYNSPYHRVTDVAVVQNEGAGAFYRSYTTQLTMNVPFQAI 241
Db 119 lthdaaisptdvikrgmqmynspytvsvscvrdyikregfkafyragtqlvmnlpqyti 178
QY 242 HPMTEYFLOEHFNPPRRNPSSHVLSGACAGAVAAATPLDVCCKTLTLNTOESLALNSHI 301
Db 179 hlttyeqnkmnlerkynppvhmaagaagaaacaaavttpldviktllntqe----- 230
QY 302 TGHITGMASAFRTYQVGGVTAYFRGVOARVIYQIPSTAIANSVYEFFKYL 353
Db 231 tgltrgmieasrkiymagpifgrgtarvlmsypataicawstwyeffkyl 282

RESULT 14
AAM80023
ID AAM80023 standard; Protein; 199 AA.
XX
AC AAM80023;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3669.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0654936.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI: 2001-476283/51.
DR N-PSDB; AAK53156.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
PS Claim 20; Page 409-410; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Notes: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 199 AA;

Query Match 38.1%; Score 739.5; DB 22; Length 199;
Best Local Similarity 71.6%; Pred. No. 1e-66;
Matches 136; Conservative 27; Mismatches 26; Indels 1; Gaps 1;

QY 174 GAAGCVATLLHDAAMNPAEYVKORMQMYNSPYHRVTDVAVVQNEGAGAFYRSYTTQLT 233
Db 8 giagsmatllhdavmmpaevvkvqrgmynsqhirsaiscirtvwtreglgafyrsyttqlt 67
QY 234 MNVPFQAIHFMTEYFLOEHFNPPRRNPSSHVLSGACAGAVAAATPLDVCCKTLTLNTOE 293
Db 68 mnvpfqsibfityeflqeqvnphtynpqshisgglagalaaaattpldvcktlntqe 127
QY 294 SLALN-SHITGHITGMASAFRTYQVGGVTAYFRGVOARVIYQIPSTAIANSVYEFFKYL 352
Db 128 nvalslanisgrlsgmanaftrvyqnglagfykqarvlygmpstaiswsavyeffkyl 187
QY 353 ITRKOEWRA 362
Db 188 ltkrqlenra 197

RESULT 15
AAM93909
ID AAM93909 standard; Protein; 187 AA.
XX
AC AAM93909;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 4059.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:37:18 ; Search time 48.11 Seconds
(without alignments)
184.804 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGVGAGGPAAGPG.....VYEFKYLITKROENRAGK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	511	26.3	328	4	US-09-068-140A-15
2	510	26.3	289	4	US-09-068-140A-10
3	293	15.1	351	2	US-08-933-750C-19
4	293	15.1	351	4	US-09-234-613-19
5	265	13.7	447	4	US-09-160-119-4
6	265	13.7	674	4	US-09-160-119-2
7	259	13.4	320	2	US-08-933-750C-12
8	259	13.4	320	4	US-09-234-613-12
9	255.5	13.2	469	4	US-09-188-930-339
10	229.5	11.8	311	2	US-08-775-009-32
11	221	11.4	311	2	US-08-775-009-33
12	201.5	10.4	432	2	US-08-937-466-4
13	201.5	10.4	432	2	US-09-172-528-4
14	201.5	10.4	432	3	US-09-318-199-4
15	201.5	10.4	432	4	US-09-503-579-4
16	198	10.2	306	5	PCT-US94-09799-1
17	197.5	10.2	308	2	US-08-937-466-2
18	197.5	10.2	308	2	US-09-172-528-2
19	197.5	10.2	308	3	US-09-318-199-2
20	197.5	10.2	308	4	US-09-503-579-2
21	186	9.6	312	4	US-09-142-565-2
22	183.5	9.5	307	2	US-08-807-861A-56
23	183.5	9.5	307	3	US-09-210-681-56
24	183.5	9.5	307	3	US-08-946-719A-56
25	181	9.3	256	2	US-08-937-466-6
26	181	9.3	256	2	US-09-172-528-6
27	181	9.3	256	3	US-09-318-199-6

28 181 9.3 256 4 US-09-503-579-6
29 171.5 8.8 309 1 US-08-518-878B-51
30 171.5 8.8 309 2 US-08-807-861A-51
31 171.5 8.8 309 2 US-08-470-868A-51
32 171.5 8.8 309 3 US-09-210-681-51
33 171.5 8.8 309 3 US-08-946-719A-51
34 168.5 8.7 299 1 US-08-518-878B-56
35 168.5 8.7 299 2 US-08-470-868A-56
36 139.5 7.2 303 1 US-08-518-878B-37
37 139.5 7.2 303 1 US-08-294-522B-36
38 139.5 7.2 303 2 US-08-807-861A-37
39 139.5 7.2 303 2 US-08-470-868A-37
40 139.5 7.2 303 3 US-09-210-681-37
41 139.5 7.2 303 3 US-08-946-719A-37
42 138.5 7.1 298 3 US-08-961-871-10
43 121.5 6.3 312 4 US-09-188-930-142
44 101.5 5.2 1155 1 US-08-094-948A-29
45 101.5 5.2 1155 5 PCT-US96-09319-29

ALIGNMENTS

RESULT 1
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 328 amino acids
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-15

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-08-933-750C-19

Query Match 15.1%; Score 293; DB 2; Length 351;
Best Local Similarity 25.1%; Pred. No. 7.7e-22;
Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;
QY 67 PAGATVTTTHWAGAVAGILEHCVMYPIDCVKTRMOSLOPDPA----- 109
DB 6 PAGISPLQOMVAGSGTCAVVTSLFMTPLDVVKVRLQSQRSPMASLMPSSRLWSLSYTKWK 65
QY 110 -----RYRNVLEALWRIIRTEG-----LWRPMLGLNVT 137
DB 66 CLLYCNGVLEPLVLCPCNGARCATWFDPTFRFTGMDAFVKIVRHEGRTLW---SGLPAT 122
QY 138 ATGACPAHALYFACYEKLLKTL-----SDVIHPGNSHIANGAACCVATLLHDAAMNPA 191
DB 123 LVMTVPATAIYFTAYDQLKAFCLGRALTSDLAP-----MVAGALARLGTVTVISPL 174
QY 192 EVVKORMQYNSPYHRVTDVCRVAVQWNEGAFYRSYTTQLTMNVPFOAIHFMTEYFLOE 251
DB 175 ELMRTKLAQHVSRELGCNVRTAVAQGWRSLSLWLGWGTALRDVFFSALYFNFELVKS 234
QY 252 HFN---PQRRYNPSSHVLGACAGAVAAATPLDVCKT-----LLNTOESLALNSHITGH 304
DB 235 WLNGLRPKDQTSVGMSEFVAGGISGTVAAVLTLPFDVVKTRQVQALGAMAEAVRVNPL--H 291
QY 305 ITGMAAFRTVYQGVGTAYFRGVARVIYQIPSTAIAMSVYEFKYLITKQEE 359
DB 292 VDSWTLRLRRRAESGTGKLFAGFLPRIKAAPSCAIMISTYEFKGFQRLNQD 346

RESULT 4
US-09-234-613-19
Sequence 19, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Nell C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-09-234-613-19
Query Match 15.1%; Score 293; DB 4; Length 351;
Best Local Similarity 25.1%; Pred. No. 7.7e-22;
Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;
QY 67 PAGATVTTTHWAGAVAGILEHCVMYPIDCVKTRMOSLOPDPA----- 109
DB 6 PAGISPLQOMVAGSGTCAVVTSLFMTPLDVVKVRLQSQRSPMASLMPSSRLWSLSYTKWK 65
QY 110 -----RYRNVLEALWRIIRTEG-----LWRPMLGLNVT 137
DB 66 CLLYCNGVLEPLVLCPCNGARCATWFDPTFRFTGMDAFVKIVRHEGRTLW---SGLPAT 122
QY 138 ATGACPAHALYFACYEKLLKTL-----SDVIHPGNSHIANGAACCVATLLHDAAMNPA 191
DB 123 LVMTVPATAIYFTAYDQLKAFCLGRALTSDLAP-----MVAGALARLGTVTVISPL 174
QY 192 EVVKORMQYNSPYHRVTDVCRVAVQWNEGAFYRSYTTQLTMNVPFOAIHFMTEYFLOE 251
DB 175 ELMRTKLAQHVSRELGCNVRTAVAQGWRSLSLWLGWGTALRDVFFSALYFNFELVKS 234
QY 252 HFN---PQRRYNPSSHVLGACAGAVAAATPLDVCKT-----LLNTOESLALNSHITGH 304
DB 235 WLNGLRPKDQTSVGMSEFVAGGISGTVAAVLTLPFDVVKTRQVQALGAMAEAVRVNPL--H 291
QY 305 ITGMAAFRTVYQGVGTAYFRGVARVIYQIPSTAIAMSVYEFKYLITKQEE 359
DB 292 VDSWTLRLRRRAESGTGKLFAGFLPRIKAAPSCAIMISTYEFKGFQRLNQD 346

RESULT 5
US-09-160-119-4
Sequence 4, Application US/09160119A
Patent No. 6316219
GENERAL INFORMATION:
APPLICANT: KRIEF, STEPHANE
APPLICANT: SOUCHET, MICHEL
APPLICANT: BRIL, ANTOINE
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GH-30985
CURRENT APPLICATION NUMBER: US/09/160,119A
CURRENT FILING DATE: 1998-09-24
EARLIER APPLICATION NUMBER: EP 97402511.6
EARLIER FILING DATE: 1997-10-23
EARLIER APPLICATION NUMBER: EP 98401655.0
EARLIER FILING DATE: 1998-07-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0

[illegible]

Query Match	13.4%;	Score 259;	DB 4;	Length 320;	
Best Local Similarity	24.7%;	Pred. No. 2.1e-18;			
Matches	78;	Conservative 62;	Mismatches 134;	Indels 42;	Gaps

QY	63	YEALPAGATVTHM--VAGAVAGILEHCVMYPIDCVKTRMOSLQ-----PDPAARYRN	114
DB	4	YDPKPDGRNNTKFOVAVAGSVSGUVTTRALISPPDVIKIRFQ-LQHERLSRSDPSAKYHGI	62
QY	115	LEALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKLTSDVIHPGG---	167
DB	63	LQASRQILQEBGPTAFKNGHVPQAQILSIGYGAQVQLSF-----EMLTELVHRGSVIDA	115
QY	168	---NSH-IANGAACCVATILLHDAAMNPAEVVVKORMQMYNSP--YHRVTCVRAVWQEGA	221
DB	116	REFSVHFVCGGLAACMATL---TVHPVDVLTRFAAQGEPKYNTLRLHVAVGTMRYSEGP	171
QY	222	GAFYSRYTTLTMNVPFOATHFMYEFLQBFHN-----PQRRYNPSSHVLSGACAGAVAA	276
DB	172	QVFKYGLAPTLIAITFPYAGLQFSCYSLEKLHYKWAIPAEGKKNNENLONLLCGSGAVISK	231
QY	277	AATTPDLVDCKTLLNT---QESLALNSHITGHITGMASAFRTVYQVGGVTAYFRGVQARVI	333
DB	232	TLTYPLDLFKRLQVGGFEHARAAGQVRRY-KGLMDCAKQVLQKEGALGFFKGLSPSLL	290
QY	334	YQIPSTAIANSVYEFF 349	
DB	291	KAALSTGFMFFSYEFF 306	

RESULT 9

US-09-188-930-339

Sequence 339, Application US/09188930A

Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: Compositions And Methods For Their Use

FILE REFERENCE: 11000.1011c1

CURRENT APPLICATION NUMBER: US/09/188,930A

CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 339

LENGTH: 469

TYPE: PRT

ORGANISM: Mouse

US-09-188-930-339

Query Match	13.2%;	Score 255.5;	DB 4;	Length 469;	
Best Local Similarity	28.6%;	Pred. No. 8.4e-18;			
Matches	84;	Conservative 51;	Mismatches 120;	Indels 39;	Gaps

QY	75	HMVAGAVAGILEHCVMYPIDCVKTRMOSLQPDPAARYN--VLEALWRIIRTEG---LNR	129
DB	189	HLVAGGAGAVSRTCTAPLDRLLKVLMOV-----HASRSNNMCIVGGFTQMIREGGAKSLWR	244
QY	130	PMRGLNVTATGAGPAHALYFACYEKLKTL-SDV---IHPGNSHIANGAACCVATILLH	184
DB	245	G-NGINYLK--IAPESAIFKPMAYEQMKRLVGSQDETLLRIH---ERLVAAGLAGAIA----	294
QY	185	DAAMNPAEVVKORMQYNS-PYHRVTCVRAVWQEGAFYRSYTTTOLTMNVPFOAIFH	243
DB	295	QSSIIYPMEVLLKTRMALRKTGOYSGMLDCARRILAKEGVAFYKGYIPNMLGIIPYAGIDL	354
QY	244	MTYEFQLQHFNPORRY-----NFSHVL--SGACAGAVAAAAATTPDLVDCKTLLNTQESLA	296
DB	355	AVYETLKNTW--LORYAVNSADPGVFVLLACGRTISSTCGOLASVPLALVTRTRMOAASIE	412


```

QY 288 LLNTQESLALNSHTTGHTGMSAFRTVIQV-----GVTRAYFRGVOARVIYQIPSTA 340
      :      |      |      |      |      |      |      |      |      |
Db 247 RMQLE-----AHKYRNTLDCGVQILKNEGPRAYKGTVPRLGRVCLDVA 291
      :      |      |      |      |      |      |      |      |      |
QY 341 IANSVYEFFKYLITK 355
      |      |      |      |      |      |      |      |      |
Db 292 IVFYIDEVVKLLNK 306
      :      |      |      |      |      |      |      |      |

```

```

RESULT 12
US-08-937-466-4
; Sequence 4, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36, 627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-937-466-4

```

Query Match	10.4%	Score	201.5	DB 2	Length	432			
Best Local Similarity	25.3%	Presd. No.	2.6e-12						
Matches	78	Conservative	44	Mismatches	139	Indels	47	Gaps	12
Qy	71	TVTHMVAGAVAGLEHCVMYPIDCVKTRMOSLPDPAA---RYRNVLLEALWRIIRTEGL	127						
Db	13	TVVKFLGAGTAACFAD-LLTFPLDTAKVRLQIQENPGCAQSQVQYRGVLGTLTWVRTECV	71						
Qy	128	WRPNRGLNVATGAGPAHALYFA----CYEKLKTTLSDVTHPGGNSHIANG---AAGCV	179						
Db	72	RSPYSGL-----VAGLHRQMSFASIRIGLYDSVQKQFYT----PKGADHSSVAIRLAGCT	122						
Qy	180	ATLLHDAAMPNPAEYVQKRMQY-----NSPYHRVTDVCRVAVQNEGACAFAYSYYTQL	232						
Db	123	TGAMAVTCAQPTDVVVKRFQAMIRLGTGGERKRYGRTMDAYRTIAREGVRGLWKGTWENI	182						
Qy	233	TMNVPFOAIHFWIYEFLOEHNPORRY--NPSSHVLSGACAGAVAAATTPDLVCKTLLN	290						
Db	183	TRNAIVNCAEMVTYDIKEXLLESHLFTDNPFCHEVSAFGFCATVVASPDVVVKRY-	241						
Qy	291	TQESIALNSHTGHITGMSAFRIVYQGVGTAYFRGVQARVVIQIPS--TAIAWSVVEF	348						

```

Db      242   -----NMAPIGRYRSPHLHCLMKMVAO-EGPTAFYKGF-----VPSFURLGAWNMMEF 287
Qy      349   FKYLITKR 356
          |           ||
Db      288   VTYEQLRK 295

RESULT 13
US-09-172-528-4
; Sequence 4, Application US/09172528
; Patent NO. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
```

```

1 CLASSIFICATION:
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: 08/937,466
4
5 FILING DATE:
6
7 ATTORNEY/AGENT INFORMATION:
8
9 NAME: OSMAN, RICHARD A
10 REGISTRATION NUMBER: 36,627
11 REFERENCE/DOCKET NUMBER: T97-0
12
13 TELECOMMUNICATION INFORMATION:
14
15 TELEPHONE: (650) 343-4341
16 TELEFAX: (650) 343-4342
17
18 INFORMATION FOR SEQ ID NO: 4:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 432 amino acids
21 TYPE: amino acid
22 STRANDEDNESS: single
23 TOPOLOGY: linear
24
25 MOLECULE TYPE: peptide
26
27 US-09-172-528-4

```

Query Match	10.4%	Score	201.5	DB	2	Length	432
Best Local Similarity	25.3%	Pred. No.	2.6	12			
Matches	78	Conservative	44	Mismatches	139	Indels	47
Gaps	12						
Qy	71	TVTTHMVAGAVAGTLEHCVMYPIDCVKTRMGSLOPDPA	---	RYRNVLEALWRIIRTEGL	127		
			:		:		:
			:		:		:
Db	13	TVKFLGAGTAACFAD-LLTFPLOTAKVRLQIQOENPCAO	SVQYRGVLGTILTWRTESG	71			
			:		:		:
Qy	128	WRPMRGLNVWTATGAPAHALYFA-----CYEKLKKT	ISDVTHPGGNSHIANG-----AAGCV	179			
			:		:		:
Db	72	RSPYSGL-----VAGLHRQMSFASIRIGLYDSVQKQFY	T-----PKGADHSSVAIRILAGCT	122			
			:		:		:
Qy	180	ATLLHDAAMNPAEYVKORMQVY-----NSPIHRVTV	DCVRAVWQNEGACAFYRSYTTQL	232			
			:		:		:
Db	123	TGMAVYCAQPTDVVKVRFQAMIRLGTGGEKRYKGTMD	AYTIAREECVRGKMGKWPNI	182			
			:		:		:
Qy	233	TMNYPFQAIHPMTYEFLOEHPNQRYY--NPSSHVLS	GACAGAVAAAAATTPLDVCKTLLN	290			
			:		:		:
Db	183	TRNAIVNCAEMVTYDIKEKLESHLFDNPPCHFVS	AFAGFGCATVVASPDVVVKRY--	241			
			:		:		:

```

Db 183 TRNAIVNCAEMVTYDIIKEKLLSHLETDNFPCHFVSAGAGCATVVASPVDVVKTRY - 241
Qy 291 TQESLALNSHITGHITGMSAFRTYVQGVGTAYFRGVQARVIYIQIPS--TAIAWSYIEF 348
Db 242 -----NNAPLGRYRSPHLCHMLKVAQ-EGTAFYKGF-----VPSEURLGAWNVMME 287
Qy 349 FKYLIITKR 356
Db 288 VTYEQLKR 295

RESULT 15
US-09-503-579-4
; Sequence 4, Application US/09503579
; Patent No. 6248561
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-503-579-4

```

Query Match	10.4%;	Score 201.5;	DB 4;	Length 432;
Best Local Similarity	25.3%;	Pred. No. 2.6e-12;		
Matches 78: Conservative	44;	Mismatches 139;	Indels 47;	Gaps 12;

Qy	71	TVTTHMVAGAVAGATLEHCVMYPPDCVKTROSLOPDAA---RYRNVEALWRIETEG	127
Db	13	TVVFLGAGTAACTAD-LLTFFLTAKVRIQIOGENPGAQSVQIRGVGLTILTWRIE	71
Qy	128	WRPGRGNVTATGAGPAHALYFA-----CYEKLKLTLSVDVHPGNSHIANG---AAGCV	179
Db	72	RSPLYSL-----VAGLHRQMSFASIRIGLYDSVKQFYT---PKGADHSVAIRILAGCT	132
Qy	180	ATLHDAAMNPAEVVKORMQMY-----NSPHRVTVDCVRVWQNEGAGAFYSTTOL	232
Db	123	TGMAVTCAPDTPDVVKYRFOAMIRLGTGGBERKRGTMDAYRTTAREEGVRLKMGKTPNI	182

QY 233 TMNVPQAIHFWTYEFLQEHFNPRRY--NPSSHVLGACAGAVAAAAATPLDVCKTLN 290
| | :||: :|| : | : | | :||: ||
Db 183 TRNAIVNCAEMVYDIKEKLLSHLFTDNFPCHFVSATGAGCATVWASPDVVKTRY- 241
QY 291 TOESLALNSHITGHTGMSAFRTVYQGGVTAYFRGVOARVIYQIPS--TAIAWSVYEF 348
:|: : : : | | ||::| :|| :||: |
Db 242 -----MNAPLGRYRSPHLCMLKMWQA-EGPTAFYKGF-----VPSFLRLGAWNVMF 287
QY 349 FKYLITKR 356
| |
Db 288 VTYEQLKR 295

Search completed: August 27, 2002, 03:58:49
Job time: 4891 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 02:35:14 ; Search time 80.04 Seconds
(without alignments)
436.988 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGGAGGAGGAGG.....VYEFKYLTKRQEWRAKG 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	799	41.2	312	2	T26089		hypothetical prote
2	601	31.0	303	2	T39149		probable RNA splic
3	598.5	28.8	304	2	S13533		mtRNA splice defec
4	555.5	28.6	314	2	S51179		mtRNA splice defec
5	521	26.9	781	2	A86205		hypothetical prote
6	507	26.1	331	2	T00582		probable mitochond
7	313.5	16.2	368	2	S54524		hypothetical prote
8	297	15.3	367	2	G36770		hypothetical prote
9	295	15.2	328	2	T19322		hypothetical prote
10	293	15.1	351	2	T43493		hypothetical prote
11	282.5	14.6	902	1	S54495		probable carrier p
12	274	14.1	269	2	T02090		hypothetical prote
13	274	14.0	330	2	T09362		hypothetical prote
14	272.5	14.0	307	2	S60949		probable phosphate
15	266.5	13.7	436	2	J01459		Btl protein precu
16	265.5	13.7	284	2	S45458		PET8 protein - yea
17	250.5	12.9	322	2	S57116		probable carrier p
18	250	12.9	338	2	T40968		probable mitochond
19	249.5	12.9	384	1	S44092		probable carrier p
20	245.5	12.7	309	2	T48156		hypothetical prote
21	240.5	12.4	475	2	T50686		peroxisomal Ca-dep
22	240	12.4	392	2	T05350		adenylate transloc
23	239.5	12.3	352	2	T01729		mitochondrial solu
24	235	12.1	326	2	S57544		probable membrane
25	234.5	12.1	335	2	T50393		probable mitochond
26	230.5	11.9	345	2	T37576		probable mitochond
27	229.5	11.8	311	2	G01789		citrate transporte
28	228.5	11.8	336	2	T01839		hypothetical prote
29	228	11.8	373	2	S48451		probable membrane

30	226	11.6	650	2	T32897	hypothetical prote
31	225.5	11.6	366	2	S64589	hypothetical prote
32	225	11.6	335	2	S50453	hypothetical prote
33	224.5	11.6	588	2	T22688	hypothetical prote
34	224	11.5	347	2	T50990	hypothetical prote
35	223	11.5	479	2	T49871	peroxisomal Ca-dep
36	222.5	11.5	302	2	T38879	probable mitochond
37	221	11.4	296	2	B96830	hypothetical prote
38	221	11.4	311	2	A46595	tricarboxylate tra
39	220	11.3	277	2	T40033	probable mitochond
40	219	11.3	271	2	S62485	probable mitochond
41	219	11.3	312	2	H89567	protein K11H3.3 [l
42	217.5	11.2	307	2	S67662	hypothetical prote
43	217	11.2	330	2	T29640	mitochondrial carr
44	216.5	11.2	349	2	B96753	hypothetical prote
45	215	11.1	702	2	T16533	hypothetical prote

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26089

R:Swinsburne, J.; Ainscough, R.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z20149

A:Accession: T26089

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-312 <WIL>

A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN000020; CESP:W02B12.9

A:Experimental source: clone W02B12

C:Genetics:

A:Gene: CESP:W02B12.9

A:Map position: 2

A:introns: 18/3; 251/3; 286/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 41.2%; Score 799; DB 2; Length 312;

Best Local Similarity 49.5%; Pred. No. 4.7e-56;

Matches 159; Conservative 54; Mismatches 82; Indels 26; Gaps 5;

QY 42 GGGEAGACRPVVRQDPDSDGPDYEALPAGATVTHMVAGAVAGILEHCVNYPIDCVKTRMQ 101

Db 3 GGGE-----DEVELPT-HSPVHLTAGALAGAVEHCVNPFDSVKTRMQ 46

QY 102 SIQPPAARYRNVLALWRIIRTEGLWRPRLNVTATGAGFAHALYFACYEKLKLTSD 161

Db 47 SLPCPCETKPTPVHSLMSIVKREGWLRPLRGVNAAGSMPAHALYFTVYKMKGYLTG 106

QY 162 VTHPGNSH-IANGAGCVATLLHDAAMNPAEVRKQMNSPYHRVTDVCRVAVWNEG 220

Db 107 --NSAGHSNTLAYAGSVVATLIHDAIMPAEVRKQMNSPYHRVTDVCRVAVWNEG 164

QY 221 AGAFYSYTTQLTMNVPPQAIHFMTYEFLOEHFNQRRYNPSHVLGACAVAAAAAT 280

Db 165 VAAFYRSTQLAMNVPPQAIHFMTYEFLOEHFNQRRYNPSHVLGACAVAAAAAT 224

QY 281 PLDVCKTLTQESALNS-----HTGHTGTGMAAPRTVYQGVGTAYPRGQARVI 333

Db 225 PMDCVKTVLNTQAAEADPANRRIELQARYRYRGISDAVRTIYSORGLSGCGLQARVI 284

QY 334 YQIPSTATAWSVYEFKYLIT 354

Db 285 FOVPATALSWSVYELFKFMLS 305

RESULT 2

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-328 <WIL>

A:Cross-references: EMBL:Z46787; PIDN:CAA86739.1; GSPDB:GN00021; CESP:C16C10.1

A:Experimental source: clone C16C10

C:Genetics:

A:Gene: CESP:C16C10.1

A:Map position: 3

A:Introns: 22/1; 125/2; 179/1; 298/1

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match

Best Local Similarity 15.2%; Score 295; DB 2; Length 328;

Matches 101; Conservative 48; Mismatches 146; Indels 68; Gaps 14;

QY 22 SPGESALLDGNLQGVGRGAGGERACRPP-----VRQDPDSGPDYEALPAGATVTTTH 76

DB 3 SPCE-----GKTTNCSGAPSSCVTPLDVVKIRLQQQTRP-----FPRGECFYH- 48

QY 77 VAGAVAGILEHCVMYPIDCVKTRMQSLQP-----DPAARYRNVLALWRIIRTEGLWRPMP 132

DB 49 -----NGLMEHV-----CVSCEVR--KPEWYQRPNGRGTADAIVKIRARHEGIRSLWS 95

QY 133 GLNVATGAGPAHALYFACYEKKKLTSDV1-----HPGNSHIANGAAGCVAT 181

DB 96 GLSPTWVWLPATVYFTYDNLVWLKKKCCRRAFSPKWTTPDWS--AAVAGIVAR 153

QY 182 LHDAAMPNPAEVVKORMQMYNPFYHRVTCVRAVWQNEGAGAFYRSYTTQLTNNVFPQAI 241

DB 154 TIATVWVSPFEMIRTKMQSKRTYHIGHLVRSSMATKGISFYLGWTPTMLRDIPFSGI 213

QY 242 HPMTEYFLEQHEPNPORYNPSSH-----VLSGACAGAVAAATTPLDVCKTLLNTQESLA 296

DB 214 YWAGVLDLEKT--NLQRQGPDPNPFVSVFSGAAGVVASIIFHPDVIKT--NCQ---- 265

QY 297 LNSHITGHTGMASAFRTV-----YQGVGTAYFRGVQARVYIQIPSTAIWSVYEFKYL 352

DB 266 --IRIGSIDDMNKSTITTVIKDMYHSGISAFSSGLVPLRVKVSPPSCALMISFYEFKFL 323

QY 353 ITK 355

DB 324 FOX 326

RESULT 10

T43493

hypothetical protein DKF2p434C119.1 - human

C:Species: Homo sapiens (man)

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C:Accession: T43493

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22516

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-351 <AA>

A:Cross-references: EMBL:AL133584

A:Experimental source: adult testis; clone DKF2p434C119

C:Genetics:

A:Note: DKF2p434C119.1

Query Match

Best Local Similarity 15.1%; Score 293; DB 2; Length 351;

Matches 89; Conservative 50; Mismatches 140; Indels 76; Gaps 8;

QY 67 PAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPA----- 109

DB 6 PAGISPLQOMVASGTGAVVTSLSFTPLDVVKVRLQSQRPMSASELMPSSRLMSLSYTKWK 65

QY 110 -----RYRNVLEALWRIIRTEG---LWRPMPGLNVT 137

DB 66 CLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVREGTRTLM-----SGLPAT 122

QY 138 ATGAGPAHALYFACYEKKKLT-----SDVHPGNSHIANGAAGCVATLLHDAAMPNPA 191

DB 123 LVMTVPATAIYFTAYDQLKAFLCGRALTSOLYAP-----MVAGALARGLGTVTVISPL 174

QY 192 EVVKORMQMYNSPYHRVTCVRAVWQNEGAGAFYRSYTTQLTNNVFPQAIHFHTYEFLOE 251

DB 175 ELMRTKLQAOHVSYSRELGACVTRTAVAGGWRSLWLGWGPPTALRDVFPFSALYFNVELVS 234

QY 252 HFN---PQRRYNPSSHVLSCACAGAVAAATTPLDVCKT-----LLNTQESLALNSHITGH 304

DB 235 WLNLGRPKDOTSGVMSFVAGGISGTVAAVLTLFPDVVKTORQVALGAMAEAVRVNPL--H 291

QY 305 ITGMAFAFTVYQVGVGTAYFRGVQARVYIQIPSTAIWSVYEFFKYLITKROE 359

DB 292 VDSWTLLLRIRAESGTKGLFAGFLPRIIKAAPSCAIMISTYEFKGSFFQRLNOD 346

RESULT 11

S54495

probable carrier protein YPR021c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YP9367.01c

C:Species: Saccharomyces cerevisiae

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 22-Oct-1999

C:Accession: S54495

R:Badcock, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54059

A:Accession: S54495

A:Molecule type: DNA

A:Residues: 1-902 <BAD>

A:Cross-references: EMBL:Z49274; NID:9809585; PIDN:CAA89275.1; PID:9809586; GSPDB:GNO

A:Experimental source: strain AB972

C:Genetics:

A:Gene: MIPS:YPR021c

A:Map position: 16R

C:Superfamily: probable carrier protein YPR021c; ADP,ATP carrier protein repeat homol

C:Keywords: duplication; transmembrane protein

F:527-615/Domain: ADP,ATP carrier protein repeat homology <ACP1>

F:621-711/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:724-814/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match

Best Local Similarity 14.6%; Score 282.5; DB 1; Length 902;

Matches 86; Conservative 45; Mismatches 129; Indels 33; Gaps 8;

QY 79 GAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWRIIRTEGLWRPMPGLNVTA 138

DB 537 GSIAGCIGATVYPIIDFIFIKRMAQR--SLAQYKNSIDCLLKIIISREGIKGLYSGLPOL 594

QY 139 TGAGPAHALYFACYEKKKLTSD-----VIHPGNSHIANGAAGCVATLLHDAAMPNPAEV 193

DB 595 IGVAPEKAIKLVNDPMPNRNLTDKNGKLSLEPEIISGASAGACQVIFT-----NPLEI 647

QY 194 VKORMQMYNSPY-----HRVTCVRAVWQNEGAGAFYRSYTTQLTNNVFPQAIHFHTYEF 248

DB 648 VKIRLQV-QSDYVGENIQOANETATQIVKRLGLRGLYNGVAAACLMRDVDPFSAYFYPTAH 706

QY 249 LOE---HFNP-----QRRYNPSSHVLSCACAGAVAAATTPLDVCKTLLNTQESLALNSH 300

DB 707 LKKDLDFDPNDKTKNRNLTWELLTAGAIGAMPAAFLTPDFDVIKRQIDPRKGETKY 766

QY 301 ITGHTIGMASAFRTVYQGVGTAYFRGVQARVYIQIPSTAIWSVYEFFKYL 353

DB 767 ----NGIFHAIRTIKESFRSFFKGGGARVLRSSPQFGFTLAAAYELFKGFI 814

RESULT 12

T20290

hypothetical protein D1046.3 - Caenorhabditis elegans

[illegible]

RESULT 15

JQ1459

Btl protein precursor - maize
C:Species: Zea mays (maize)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Sep-1999
C:Accession: JQ1459
R:Sullivan, T.D.; Strelow, L.I.; Illingworth, C.A.; Phillips, R.L.; Nelson Jr., O.E.
Plant Cell 3, 1337-1348, 1991
A:Title: Analysis of maize brittle-1 alleles and a defective suppressor-mutator-induced
A:Reference number: JQ1459; MUID:93005685
A:Accession: JQ1459
A:Molecule type: mRNA
A:Residues: 1-436 <SU>
A:Cross-references: GB:W79333; NID:g168425; PIDN:AAA33438.1; PID:g168426
C:Comment: This protein acts as an adenylate translocator in amyloplasts.
C:Genetics:
A:Introns: 206/3; 260/3
C:Superfamily: Btl protein; ADP,ATP carrier protein repeat homology
C:Keywords: chloroplast; duplication; transmembrane protein
F:/15/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:/76-436/Product: Btl protein #status predicted <MAT>
F:/131-217/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:/226-312/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:/229-247/Domain: transmembrane #status predicted <TM1>
F:/323-413/Domain: ADP,ATP carrier protein repeat homology <ACP3>
F:/327-347/Domain: transmembrane #status predicted <TM2>

Query Match	13.7%	Score 266.5;	DB 2;	Length 436;
Best Local Similarity	25.9%;	Pred. No. 1.5e-13;		
Matches	98;	Conservative	57;	Mismatches 138; Indels 85; Gaps 16;
QY	19	PG-RSPGESALLDGLQRGVGRGAGGEGACACR-----PPVRQDPDSGDGYEALPAGATV	72	
Db	90	PGSRPPCRR-----GRGSEEEAEGRRHEAAAGRSEPEEGQGDQRQAPAA--	136	
QY	73	TTTHWAGAVAGAILHCVMYPIDCVKTRMQ--SLQDPAPARYNNVLEALWRIIRTEGLWRP	130	
Db	137	--RLVSGAIIAGAVSRTEVAPLETIRTHLWGSIGVDSMA---GVFQ--W-IMQEGWTGL	188	
QY	131	MRLGNVTATGAGPAHAIFYCYEKLKLTSDVHPGNSH-----TANGAAGCVAT	181	
Db	189	FRGNVAVNLVRVAPSKAIEHTYTDYAKFKLT----PKGDEPPKIDPTPLVAGALAGFAST	244	
QY	182	LLHDAANPAEVYKQRMQNTSPYHRVTDVRAVWQNEGAGAFYRSTTQLTMMVFPQAI	241	
Db	245	L-----CTYPMELIKTRVTIEKDVYDNVAHAFVKILTRDEGSPSELYRGLTSPSLIGVVVPAAC	300	
QY	242	HEMTYEFLOEHF-----NQRRYNPSSHVLSGACAGAVAAAATPLDVCCKTLNTOESL	295	
Db	301	NFYAYETLKLYRRATGRREPGADVGPVATLLIGSAGAIASSATFPFLVARQKMQ-----	355	
QY	296	ALNSHITGHITGMASAFRTVYQ-----VGGVTAFRGVOARVIYQIPSTAI	341	
Db	356	-----VGAVGG-----ROYQNVLHAIYCIILKKEGAGGL---YRGLGPSICIKLMPAAGI	401	
QY	342	AWSWYEFFKYLITKROEE	359	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:56:08 ; Search time 37.5 Seconds

(without alignments)
375.838 Million cell updates/sec

Title: US-09-870-113-2

Perfect score: 1940

Sequence: 1 MELEGRGAGGAGGPAAGPG.....VYEFKYLTIKROENRAGK 364

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558.5	28.8	304	1 MRS4_YEAST	P23500 saccharomyc
2	555.5	28.6	314	1 MRS3_YEAST	P10566 saccharomyc
3	313.5	16.2	368	1 Y039_YEAST	Q03829 saccharomyc
4	295	15.2	328	1 Y051_CAEEL	Q09461 caenorhabdi
5	289	14.9	359	1 CG69_HUMAN	Q9bz14 homo sapien
6	272.5	14.0	307	1 ODC2_YEAST	Q92927 saccharomyc
7	268	13.8	676	1 CMC2_MOUSE	Q9gxx4 mus musculu
8	265.5	13.7	436	1 ETL_WAIZE	P29518 zea mays (m
9	265.5	13.7	284	1 PET8_YEAST	P38921 saccharomyc
10	265	13.7	675	1 CMC2_HUMAN	Q9ujs0 homo sapien
11	261.5	13.5	695	1 CMCL_DROME	Q9va73 drosophila
12	259	13.4	320	1 DNC_HUMAN	Q9hc21 homo sapien
13	250.5	12.9	322	1 SFCL_YEAST	P33303 saccharomyc
14	244.5	12.6	315	1 MFT_HUMAN	Q9h2d1 homo sapien
15	240	12.4	315	1 SA18_HUMAN	Q9h1k4 homo sapien
16	240	12.4	678	1 CMCL_HUMAN	Q75746 homo sapien
17	235	12.1	312	1 SA18_MOUSE	Q9db41 mus musculu
18	230.5	11.9	345	1 YDE9_SCHPO	Q10442 schizosacch
19	229.5	11.8	311	1 TXTP_HUMAN	P53007 homo sapien
20	228	11.8	373	1 YIA6_YEAST	P40556 saccharomyc
21	227	11.7	299	1 ODC_HUMAN	Q9bqt8 homo sapien
22	225.5	11.6	366	1 YG5F_YEAST	P53320 saccharomyc
23	225	11.6	335	1 YEA6_YEAST	P39953 saccharomyc
24	224.5	11.6	311	1 TXTP_BOVIN	P79110 bos taurus
25	224.5	11.6	588	1 CMC2_CAEEL	Q20799 caenorhabdi
26	223	11.5	301	1 MCAT_HUMAN	Q43772 homo sapien
27	222.5	11.5	302	1 YDIK_SCHPO	Q10248 schizosacch
28	221	11.4	311	1 TXTP_RAT	P32089 rattus norv
29	219	11.3	271	1 YAD8_SCHPO	Q09834 schizosacch
30	219	11.3	301	1 ORTL_MOUSE	Q9wvd5 mus musculu
31	219	11.3	312	1 TXTP_CAEEL	P34519 caenorhabdi
32	215.5	11.1	301	1 MCAT_RAT	P97521 rattus norv
33	215	11.1	702	1 CMCL_CAEEL	Q21153 caenorhabdi

ALIGNMENTS

```
RESULT 1
MRS4_YEAST
ID MRS4_YEAST STANDARD; PRT; 304 AA.
AC P23500;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial RNA splicing protein MSR4.
GN MSR4 OR YKR052C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1301.
RX MEDLINE=91108815; PubMed=1703236;
RA Wiesenberg G., Link T.A., von Ahnen U., Waldherr M., Schweyen R.J.;
RT "MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast,
RL J. Mol. Biol. 217:23-37(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Viissers S., Urrestazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MRS4 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE
CC FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING
CC ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
CC THE MITOCHONDRION (POSSIBLY OF CATIONS). NOT ESSENTIAL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
CC TO YEAST MRS3; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56444; CAA39828.1; -.
CC EMBL; Z28277; CAA82130.1; -.
CC PIR; S13533; S13533.
CC SGD; S0001760; MRS4.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PROSITE; PS00215; MITOCH_CARRIER; 2.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC mRNA processing.
CC REPEAT 9 110 BLOCK I (APPROXIMATE).
CC REPEAT 111 202 BLOCK II (APPROXIMATE).
CC REPEAT 203 304 BLOCK III (APPROXIMATE).
CC TRANSMEM 23 41 POTENTIAL.
```



```
ID YM39_YEAST STANDARD; PRT; 368 AA.
AC Q03829;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier YMR166C.
GN YMR166C OR YMR520.15C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Hunt S., Bowman S., Barrall B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 249705; CAA89802.1; -
DR SGB; S0004776; YMR166C.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_car; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
SQ SEQUENCE 368 AA; 40992 MW; B583100018DF045D CRC64;
```

Query Match 16.2%; Score 313.5; DB 1; Length 368;
Best Local Similarity 29.0%; Pred. No. 1.1e-17;
Matches 88; Conservative 66; Mismatches 114; Indels 35; Gaps 11;

```
QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYNVLEALWRIIRTEGLWRPMLN 135
DB 57 VVSGIGGKIGDSAMHSLDTVKTRQG-APN-VKRYRNMSAYRTIWLBEGVRRGLGGY 114
QY 136 VTA-TGAGPAHALYFACYEKLKLTSDVIHPGNGSHIANGAGCVATLHDAAMNPAEV 194
DB 115 MAALGSPSPAIFGTGYEYKRTMEDQI--NTITHLSAGFLGDFISSFVYVPSEVL 172
QY 195 KORMQM---YNSP-----YHRVTDVRAVQNGAGAFYSYTTQLTMNVFQAIHFM 244
DB 173 KTRLOLQGRFNNPFQSGYNYNLRAIKTVKEGFRSLFFGYKATLARDLPFSALQFA 232
QY 245 TYE-FLOEHFNQRR-----YNPSSHVLSGACAGAVAAAATPLDVKCTLLNTOESLA 297
DB 233 FYEKRFLAQFKEQDGRDGLSIPNEILTGACAGGLAGIITPMDVVKTRVTOQPPSQ 292
QY 298 NS-----HIT-GHITGMAS-----AFRTVYQGVGTAYFRGVOARVIYQIPSTAIWS 344
DB 293 SNKSYSVTHPHVTNGRPAALNSISLRTVYQSEGLGFFSGVGRFVWTSVQSIMLL 352
QY 345 VYE 347
DB 353 LYQ 355
```

RESULT 4
YQ51_CAEEL

```
ID YQ51_CAEEL STANDARD; PRT; 328 AA.
AC Q09461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier C16C10.1.
GN C16C10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Lloyd C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 246787; CAA86739.1; -
DR WormPep; C16C10.1; CE01489.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_car; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Repeat; Transmembrane;
KW Transport.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 328 AA; 36743 MW; EA1E9E329A764DF6 CRC64;
```

Query Match 15.2%; Score 295; DB 1; Length 328;
Best Local Similarity 27.8%; Pred. No. 2.8e-16;
Matches 101; Conservative 48; Mismatches 146; Indels 68; Gaps 14;

```
QY 22 SPGESALLDGLQGVGRGAGGEGAGACRPP-----VRQDPDGPDPYEALPAGATVTTM 76
DB 3 SPES-----GKTTNCSGAPSSSCVTPLDVVKIRLQQQTRP----FPGCECFYH- 48
QY 77 VAGAVAGILEHCVMYPIDCVKTRMQSLQ-----DPAARYNVLEALWRIIRTEGLWRPML 132
DB 49 -----NGLMEHY-----CVSCEVR--KPCWYQRPNGFRGTADAIVKIHGIRSLWS 95
QY 133 GLNVTATGAGPAHALYFACYEKLKLTSDVI-----HPGNGSHIANGAGCVAT 181
DB 96 GLSPFMVMPATVTFYTTIDNLVWLKKKMCRRFAFSPEKTPPDWS--AAVAGIAR 153
QY 182 LLHDAAMNPAEYVKORMQYNIPYHRVTDVRAVQNGAGAFYSYTTQLTMNVFQAI 241
DB 154 TIATVTVSPPIEMIRTKMQSKRLTYHEIGHVRSRMATKGISSFLGWTPTMLRDIFSGI 213
QY 242 HMTVEFLOEHFNQRRYNPSSH-----VLSGACAGAVAAAATPLDVKCTLLNTOESLA 296
DB 214 YWAGYDLFT--NLORROGPDHNPFFVSVSAGAAVWASIFTHPPDVTKT--NCO---- 265
QY 297 LNSHTGHTGTGNASAFRTV---YQVGGVTAYFRGVOARVIYQIPSTAIWSYEFKYL 352
DB 266 --IRGGSIDDMKNSITTVIKDMYHSRGISAFSSGLVPLRVKVPSCAIMSIFYEFKFL 323
QY 353 ITK 355
DB 324 FOK 326
```

RESULT 5

CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9B274; O9Y379; Q9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
RA Pan G., Adams S.H.;
RT "Overexpression of the human 2-oxoglutarate carrier lowers
RT mitochondrial membrane potential in HEK-293 cells: contrast with the
RT unique cold-induced mitochondrial carrier CGI-69.";
RL Biochem. J. 353:369-375(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20372150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain, and Colon;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
CC are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
CC and kidney.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF317711; RAG60687.1; -
DR EMBL; AF151827; AAD34064.1; -
DR EMBL; AL133584; CAB63728.1; -
DR EMBL; AK026060; BAB15341.1; -
DR EMBL; BC001398; AAH01398.1; -
DR EMBL; BC009330; AAH09330.1; -
DR EMBL; AF119864; RAG69618.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW Alternative splicing; Polymorphism.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSPLIC 64 72 LPSSLOSTG -> W (IN ISOFORM 2).
FT VARIANT 247 247 L -> F.
FT /FTID=VAR_012756.
FT CONFLICT 266 266 G -> R (IN REF. 6).
FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;
SQ

Query Match 14.9%; Score 289; DB 1; Length 359;
Best Local Similarity 24.5%; Pred. No. 9.4e-16;
Matches 89; Conservative 50; Mismatches 140; Indels 84; Gaps 8;

QY 67 PAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDA-----109
DB 6 PAGISPLQMVASGTGAVVTSLEMTPLDVVKVRLQSORPSMASLAPSSRLWSLSYTKLP 65
QY 110 -----RYRNVLALWRIIRTEG---LWR 129
DB 66 SSIQSTGKCLLYCNGVLEPLLYLCPNGRCATWFDPTFRFTGMDAFVKIVRHGTRTLW- 124
QY 130 PMRGLNVATGAGPAHALYFACYEKLKLT-----SDVIHPGGNSHIANGAACVATLL 183
DB 125 --SGLPATLVMTVPATAIYFTAYDQLKAFCLGRALTSDLVAP-----MWAGALARLG 174
QY 184 HDAAMNPAEVVKORMQMYNSPYHRVTDCVRVWQNEGAGAFYSYTTOLTMMVYFQAIHF 243
DB 175 TVTVISPLELMRTKLOAQHVSYRELGCACVTAQAQGWGWSRLWLGWGTALRDVYFVSALYW 234
QY 244 MTVEFLQEHEN--PQRRYNPSSHVLSGACAGAVAAATPLDVCKT-----LLNTOBSLA 296
DB 235 FNYELVKSWLNGLRPKDQTSVGMFSVAGGISGVIAVLLTLPFDVVKTORVALGAEAVR 294
QY 297 LNSHITGHITGMASAFRTYVQGVGTAYFRGVOARVIYQIPSTAIAMSVYEFFKYLITKR 356
DB 295 VNPL---HVDSTWLLLRIRAESGTGLFAGFLPRIKAAPSCAIMISTVEFGKSPFQRL 351
QY 357 QEE 359
DB 352 NQD 354

RESULT 6
ODC2_YEAST ID ODC2_YEAST STANDARD; PRT; 307 AA.
AC Q99297;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxodicarboxylate carrier 2.

GN ODC2 OR YOR222W OR YOR50-12.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RA Galisson F., Dujon B.;
 RT "Sequence and analysis of a 33 kb fragment from the right arm of
 RT chromosome XV of the yeast Saccharomyces cerevisiae.";
 RL Yeast 12:877-885(1996).
 RN [2]
 RP CHARACTERIZATION.
 RX PubMed=11013234;
 RA Palmieri L., Agrimi G., Runswick M.J., Fearnley I.M., Palmieri F.,
 RA Walker J.E.;
 RT "Identification in Saccharomyces cerevisiae of two isoforms of a novel
 RT mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate.";
 RL J. Biol. Chem. 276:1916-1922(2001).
 CC -!- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
 CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
 CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
 CC citrate and malate. The main physiological role is probably to
 CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
 CC matrix to the cytosol where they are used in the biosynthesis of
 CC lysine and glutamate, respectively, and in lysine catabolism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 275130; CAA99440.1; -.
 DR EMBL; X92441; CAA63185.1; -.
 DR SGD; S0005748; ODC2.
 DR InterPro; IPR002067; Mit_carrier.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCH_CARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 SQ SEQUENCE 307 AA; 34007 MW; 4089082A64DBA97C CRC64;

Query Match 14.0%; Score 272.5; DB 1; Length 307;
 Best Local Similarity 27.2%; Pred. No. 1.6e-14;
 Matches 81; Conservative 54; Mismatches 130; Indels 33; Gaps 7;
 QY 77 VAGAVAGILEHCVMYPIDCVKTRMOSLPDPA--RNRNLEALWRIIRTEGLW 128
 DB 17 ISGAVAGISELVMTPLDVKVTRFQLEVTPTAAVGVKOVERINGVDCLKKIVKKEGS 76
 QY 129 RPRMGLNVTATGAGFAHALYFACYEKLKTKLSDVHPGNSHANGAACVATLLHDAAM 188
 DB 77 RLYRGISSPLMEAPKRAKTFACNDQYQKIFKNLNTNETTKISIAAGASAGMTAAVI 136
 QY 189 NPAEUVVKQRMQYNSPYHRVTCVRAVWQNEGAGAFYRSTYTLQTMVNFQAIHF-MTYE 247
 DB 137 VPFELIKIRMQDVKSSYLPGMDCLKTKIKNEGIMGLYKGIESTMNRNALWNGGYFQVIQ 196

QY 248 FLQEHFNPPORRYNP-----SSHLVSGACAGAVAAAATTPLDVCKTLLNTQESIALN 298
 DB 197 V-----RNSMPVAKTKQKTRNDLIAGAIIGTGTMLNTFFDVVKSHQSDV--ANS 246
 QY 299 SHITGHITGMASAFRTVYQVGVYATYRGVQARVYIQTSTAIWSVY-----EFFKYL 352
 DB 247 SAVKKNWCLPSLL-VIYREGFRALYKGFVPKVCRLAPGSGSLMLVVFVTGMNFFRDL 303
 RESULT 7
 CMC2_MOUSE
 ID CMC2_MOUSE STANDARD; PRT; 676 AA.
 AC Q9XX4; Q9DCF5; Q9CF6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Calcium-binding mitochondrial carrier protein Aralar2 (Solute carrier
 DE family 25, member 13) (Citrin).
 GN SLC25A13 OR ARALAR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20079163; PubMed=10610724;
 RA Sinasac D.S., Crackower M.A., Lee J.R., Kobayashi K., Sahaki T.,
 RA Scherer S.W., Tsui L.-C.;
 RT "Genomic structure of the adult-onset type II citrullinemia gene,
 RT SLC25A13, and cloning and expression of its mouse homologue.";
 RL Genomics 62:289-292(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo, and Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fieischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willing L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER. MAY HAVE
 CC A FUNCTION IN THE UREA CYCLE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: AT E10.45, EXPRESSED IN BRANCHIAL ARCHES AND
 CC LRTMB AND TAIL BUDS. AT E13.5, EXPRESSION IS PREDOMINANT IN
 CC EPITHELIAL STRUCTURES AND THE FOREBRAIN, KIDNEY AND LIVER.
 CC EXPRESSION IN LIVER IS MAINTAINED INTO ADULTHOOD.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF164632; AAF21426.1; -
CC EMBL; AK002829; BAB22390.1; -
CC EMBL; AK012670; BAB28397.1; -
CC MGD; MGI:1354721; Sic3sa13.
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00036; ehand; 3.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 3.
CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC KW Calcium-binding.
CC FT TRANSMEM 333 350 1 (POTENTIAL).
CC FT TRANSMEM 394 413 2 (POTENTIAL).
CC FT TRANSMEM 437 450 3 (POTENTIAL).
CC FT TRANSMEM 486 505 4 (POTENTIAL).
CC FT TRANSMEM 525 542 5 (POTENTIAL).
CC FT TRANSMEM 582 601 6 (POTENTIAL).
CC FT CA_BIND 66 77 EF-HAND 1.
CC FT CA_BIND 100 111 EF-HAND 2.
CC FT DOMAIN 139 150 ANCESTRAL CALCIUM SITE 3.
CC FT CA_BIND 171 182 EF-HAND 4.
CC FT REPEAT 207 422 1.
CC FT REPEAT 423 512 2.
CC FT REPEAT 513 615 3.
CC FT CONFLICT 312 312 MISSING (IN REF. 2; BAB28397).
CC FT CONFLICT 442 442 S -> F (IN REF. 2; BAB22390).
CC SQ SEQUENCE 676 AA; 74466 MW; 5E3952F08E5E58C5 CRC64;

Query Match 13.88; Score 268; DB 1; Length 676;
Best Local Similarity 28.28; Pred. No. 9.2e-14;
Matches 82; Conservative 47; Mismatches 122; Indels 40; Gaps 9;

QY 79 GAVAGILEHCVMYPIDCVKTRMQSLQDPDA-----ARYRNVLALWRIIRTEGLWRPWRGL 134
Dy 336 GSIAGAGATVAPYDILVKTRMQNRQSGFVGMELMKNSEDFCKKVLVRVGGFGLYRGL 395
QY 135 NVATGAGPAHALYFACYEKLKTLSDVI-----HPGNSHI-----ANGAAGCVATLLH 184
Dy 396 LPQLLGVAPEKA-----IKLTNDFVRDKFMKDGSPVLLAEIFAGGCGAGSQVIF- 446
QY 185 DAANPAEVVKORMYNSPHYRVDVCR-----AYWNEGAGAFYRSYTTQLTMNVFQA 240
Dy 447 ---TNPLEIVKIRLQVAG-----EITTPRVSAVSVRDLGFFGIYKGAACFLRDIPFSA 499
QY 241 IHFTYEFLOEHF-NPQRRYNPSSHVLSGACAGAVAAATTPLDVCKTLLNTQESLANS 299
Dy 500 IYFCYAHVKASFANEDQVSPGSLLAGAIGAPASLVTTPADVIKTRLOVARAGQTT 559
QY 300 HITGHITGMASAFRTYQVGVTAIFRGVQARVIYQIPSTAIAMSVYEFK 350
Dy 560 Y-----NGVTDCTFRKILREEGPRALMKGVAAVRSSPQGVQVTLTLLYELLQ 605

RESULT 8
BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT *Analysis of maize brittle-1 alleles and a defective
RT Suppressor-mutator-induced mutable allele.*;
RL Plant Cell 3:1337-1348(1991).
CC -|- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -|- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -|- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -|- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M79333; AAA33438.1; -
CC DR PIR; JQ1459; JQ1459.
CC DR MaizeDB; 47578; -
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 1.
CC KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
CC FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
CC FT CHAIN 76 436 BRITTLE-1 PROTEIN.
CC FT TRANSMEM 229 247 POTENTIAL.
CC FT TRANSMEM 327 347 POTENTIAL.
CC SQ SEQUENCE 436 AA; 46627 MW; 960C05F603E9DAE CRC64;

Query Match 13.7%; Score 266.5; DB 1; Length 436;
Best Local Similarity 25.9%; Pred. No. 7.1e-14;
Matches 98; Conservative 57; Mismatches 138; Indels 85; Gaps 16;

QY 19 PG-RSPGESALLDGLWQVGRGAGGEGAGACR-----PPVRQDPDSGPDYEALPAGATV 72
Dy 90 PGSRPPGRR-----GRGSEEEAEGRHHEEAAAAGSRSEPEGQGDQRPAPA-- 136
QY 73 TTHMVAGAVAGILEHCVMYPIDCVKTRMQ--SLOPDPAARYRNVLALWRIIRTEGLWRP 130
Dy 137 --RLVSGAIAAGAVSRTPVAPLETIRTHLMVSGIGVDSMA---GVFQ--W-IMQEGWTGL 188
QY 131 MRGLNVATGAGPAHALYFACYEKLKTLSDVIHPGNSH-----IANGAAGCVAT 181
Dy 189 FRGNVNVLRVAPSKAIEHTYDTAKKFLT---PKGDEPKPIPTPLVAGALAGFAST 244
QY 182 LHDHAANPAEVVKORMYNSPHYRVDVCRVAVWNEGAGAFYRSYTTQLTMNVFQA 241
Dy 245 L-----CTYPMELIKTRVTIEKVDVNAHAFVKILRDEGPSELYRGLTSLIGVVPYAA 300
QY 242 HEMTYEFLOEHF-----NPQRRYNPSSHVLSGACAGAVAAATTPLDVCKTLLNTQESL 295
Dy 301 NFAYETLKKRLYRRATGRRPGADVGVATLLIGSAGAIASSATFPLEVARKQM----- 355
QY 296 ALNSHITGHITGMASAFRTYQ-----YGGVTAYFRGVQARVIYQIPSTAI 341
Dy 356 -----VGAVGG-----RQVYQNVHLHAIYCLTKKEGAGGL---YRGLGPSCKILMPAAGI 401
QY 342 AWSVYEFKLYTKRQEE 359
Dy 402 AFMCYEACKKILVDKEDE 419

RESULT 9
PET8_YEAST STANDARD; PRT; 284 AA.
ID PET8_YEAST
```


CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF118838; AAD38501.1; -
 DR EMBL; Y17571; CAB62206.1; -
 DR EMBL; AF164530; AAF28473.1; -
 DR EMBL; AF164525; AAF28473.1; JOINED.
 DR EMBL; AF164526; AAF28473.1; JOINED.
 DR EMBL; AF164527; AAF28473.1; JOINED.
 DR EMBL; AF164528; AAF28473.1; JOINED.
 DR EMBL; AF164529; AAF28473.1; JOINED.
 DR EMBL; AC002450; AAB67049.1; -
 DR EMBL; AC002540; AAB70112.1; -
 DR MIM; 603859; -
 DR MIM; 603471; -
 DR MIM; 605814; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR002067; Mit_carrier.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR Pfam; PF00036; efhand; 3.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
 DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Calcium-binding; Polymorphism.
 FT TRANSEM 332 349 1 (POTENTIAL).
 FT TRANSEM 393 412 2 (POTENTIAL).
 FT TRANSEM 436 449 3 (POTENTIAL).
 FT TRANSEM 485 504 4 (POTENTIAL).
 FT TRANSEM 524 541 5 (POTENTIAL).
 FT TRANSEM 581 600 6 (POTENTIAL).
 FT CA_BIND 66 77 EF-HAND 1.
 FT CA_BIND 100 111 EF-HAND 2.
 FT DOMAIN 139 150 ANCESTRAL CALCIUM SITE 3.
 FT CA_BIND 171 182 EF-HAND 4.
 FT REPEAT 207 421 1.
 FT REPEAT 422 511 2.
 FT REPEAT 512 614 3.
 FT CONFLICT 231 232 EL -> VH (IN REF. 2).
 FT CONFLICT 311 311 Q -> QQ (IN REF. 4).
 FT CONFLICT 532 532 M -> T (IN REF. 2).
 SQ SEQUENCE 675 AA; AD07EDBC6C68989B CRC64;
 Query Match 13.7%; Score 265; DB 1; Length 675;
 Best Local Similarity 28.2%; Pred. No. 1.6e-13;
 Matches 82; Conservative 47; Mismatches 122; Indels 40; Gaps 9;
 QY 79 GAVAGILEHCVMYPIDCVKTRMSIQDPDA-----ARYRNVLALWRIIRTEGLWRPVRGL 134
 DB 335 GSVAGAGATATVYIDLVKTRMQNQRGTGSGVGBELTKNSFCDFKVLRYEGFGLYRGL 394
 QY 135 NVATGAGPAHALYFACYEKLKLTLSVVI-----HPGGN-----SHIANGAAGCVATLLH 184
 DB 395 LPQLLVGAPEKA-----IKLTNVDFVRDKFMHKDGSVPLAEILAGGCAGGSQVIF- 445
 QY 185 DAANPAEVVKQRMQNSPYHRVTDVCR-----AVWQNEGAGAPRYTYTQLTNVVPFQA 240
 DB 446 ---TNPLEIVKIRLQVAG-----EITTPRVVSALSVVRDLGLFGFIYKGAACFLRDIPFSA 498
 QY 241 IHEWTYEFLEHFN-PPQRYNPNSSHVLSGACAGAVAAATTPDLVCKTLLNTQESIALNS 299
 DB 499 IYPCYAHVRKASFANEDQGVSPGSLLAGAIAAGMPAASLVTPADVIKTRLQVAARAGQTT 558

QY 300 HITGHITGMASAFRTVYQVGVTAYFRGQVQARVIYQIPSTAIASVYEFFK 350
 DB 559 Y-----SGVIDCFKILREEGPKALWKAGARVFRSSPQFGVTLITVELLQ 604
 RESULT 11
 CMCL1_DROME
 ID CMCL1_DROME STANDARD: PRT; 695 AA.
 AC Q9VA73; Q9VA72; Q9VA74; Q9U5V8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calcium-binding mitochondrial carrier Aralar1.
 GN ARALAR1 OR CG2139.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Ovary;
 RX MEDLINE=20115110; PubMed=10642534;
 RA Del Arco A., Agudo M., Satrustegui J.;
 RT "Characterization of a second member of the subfamily of calcium-
 RT binding mitochondrial carriers expressed in human non-excitable
 RT tissues";
 RL Biochem. J. 345:725-732(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokskein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2195-2195(2000).
 CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
 CC SIMILARITY).


```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538421; PubMed=10978331;
RA Titus S.A., Moran R.G.;
RT "Retrovirally mediated complementation of the glyB phenotype. Cloning
RT of a human gene encoding the carrier for entry of folates into
RT mitochondria."
RL J. Biol. Chem. 275:36811-36817(2000).
CC -!- FUNCTION: Transport folate across the inner membranes of
CC mitochondria.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF283645; AAG37834.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 89 106 POTENTIAL.
FT TRANSMEM 227 243 POTENTIAL.
FT TRANSMEM 281 300 POTENTIAL.
SQ SEQUENCE 315 AA; 35388 MW; 60EBCDD61951EB6E CRC64;

Query Match 12.6%; Score 244.5; DB 1; Length 315;
Best Local Similarity 24.8%; Pred. No. 2.7e-12;
Matches 79; Conservative 54; Mismatches 142; Indels 43; Gaps 11;

QY 68 AGATVTH-----WVAGAVAGILEHCVMYPIDCVKTRMQ-----SLQDPDPAARYRNVL 117
DB 13 AWSTVFRHVRVYENLIAGVSGVLSNLAHLPLDLVKIRFAVSDGLELRP-----KYN 68

QY 118 LWRIIRTEGLMRPGLNVTATGAGPAHALYFACYEKLK--KTLSDVIHPGGNSHIANGA 175
DB 69 LTTIKWLDGLRGLYGVTPNHWAGLSGLGFFFFNAIKSKYKTEGRAHLEATEYLVSA 128

QY 176 AGCVATLLHDAAMNPAEYVVKORMQ-----XNSP---YHRVTDVVRVWQNEGAGAFYRS 227
DB 129 EAGAMTL---CITNPLWTKTFLMLQYDAVNVSPHRQYKGMFDTLVKIYKYGVRGLYKG 185

QY 228 YTTQLTMMVFPQAIFHWTFYELQ-----EHEFN--PQRRYNPSHVLVSACAGAAAAATTP 281
DB 186 FVPGI--FCTSHGALQFMAYELLKLYNOHINRLPEAQLSTVEYISVAALSIFAVAATYP 244

QY 282 LDVCKTLTNTQESLALNSHITGHTMASAFRTVYQVGGVTAYFRGQARVYQIPSTAI 341
DB 245 YQVVRARLQDQHF-----YSGVIDVTKWRKEGVGGYFKGIAPNLRVTPACCI 295

QY 342 AWSVYEFKYLITRKEE 359
DB 296 TFWYVYNSHLLDLREK 313

RESULT 15
SA18_HUMAN STANDARD; PRT; 315 AA.
ID SA18_HUMAN

```

```

AC O9H1K4;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 25, member 18.
DE SLC25A18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21275466; PubMed=11381032;
RA Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riazzi A.M.,
RA Bridgland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,
RA Nguyen T., Fang F., Fu Y., Ray L., Wu H., Shauli S., Phan S., Yao Z.,
RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,
RA McDermid H.E.;
RT "Analysis of the cat eye syndrome critical region in humans and the
RT region of conserved synteny in mice: a search for candidate genes at
RT or near the human chromosome 22 pericentromere."
RL Genome Res. 11:1053-1070(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY008285; AAG22855.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
SQ SEQUENCE 315 AA; 33848 MW; B23A9E5036671634 CRC64;

Query Match 12.4%; Score 240; DB 1; Length 315;
Best Local Similarity 25.2%; Pred. No. 6e-12;
Matches 77; Conservative 53; Mismatches 136; Indels 40; Gaps 9;

QY 71 TTTTHWAGVAGILEHCVMYPIDCVKTRMOSLQDPDPAARYRNVLALRIIRTEGLWRP 130
DB 7 SITAKLVGVAGLVGVTCVFPIDIAKRLQN--OHGRAMKGMIDCLMKTRABGFTQM 64

QY 131 MRGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGGNSHI--ANGAAGCVATLLHDAAMN 189
DB 65 YRGAANLTLVTPEKAIKLAANDFFRLL---MEDGMQRNLKWEMLAGCGAGCQVVVTC 121

QY 190 PAEVVKORMQ-----MTNSPYHRVTDVCRVWQ---NEGAGA 223
DB 122 PMEMLKIQLDAGRLAVHHQGSASAPTSRSYTTGSASTHRRPSATLIAWELLRTQGLAG 181

QY 224 FYRSYTTQLTMMVFPQAIFHWTFYELQ--HFNPQRRYNPSH--VLSSACAGAAAAATTP 281
DB 182 LYRGLGATLLRDIPFSIIFFPLFANLNGLFNLAKGAFSAHFSVGCAGVAAVATP 241

QY 282 LDVCKTLTNT--QESLALNSHITGHTMASAFRTVYQVGGVTAYFRGQARVYQIPSTA 340
DB 242 LDVLKTRITQTLKKGLEDWY-----SGITDCARKLWIOEGPSAFMKGAGCRALVIAPLFG 296

```

QY 341 IAWSVY 346
|| ||
Db 297 IAQGVY 302

Search completed: August 27, 2002, 04:16:45
Job time: 1237 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:57 ; Search time 138.54 Seconds
(without alignments)
154.737 Million cell updates/sec

Title: US-09-870-113-4
Perfect score: 1036
Sequence: 1 MELEGRGAGVGAGGAAAGPG.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	786	75.9	318	22 AAM41505	Human polypeptide
2	624	60.2	289	21 AAB50383	Human uncoupling p
3	624	49.9	299	22 AAB60113	Human transport pr
4	516.5	49.7	331	22 AAM79039	Human protein sequ
5	515	49.7	331	22 AAB60658	Human mitochondria
6	486.5	47.0	677	22 AAU29748	Novel human secret
7	477	46.0	268	22 AAM39719	Human polypeptide
8	465	44.9	366	22 AAM00938	Human bone marrow
9	440.5	42.5	155	21 AAB42966	Human OREF 2730
10	363	35.0	272	21 AAB42980	Human OREF 2744
11	295	28.5	379	22 ABB61130	Drosophila melanog

12	241.5	23.3	126	22	ABB51123	Human secreted pro
13	223	21.5	331	21	AAG22077	Arabidopsis thalia
14	223	21.5	331	21	AAG43092	Arabidopsis thalia
15	211	20.4	289	18	AAW17054	Blackcurrant Rib7
16	211	20.4	290	21	AAG22078	Arabidopsis thalia
17	211	20.4	290	21	AAG43093	Arabidopsis thalia
18	167	16.1	278	21	AAG22079	Arabidopsis thalia
19	167	16.1	278	21	AAG43094	Arabidopsis thalia
20	165.5	16.0	59	22	AAM85891	Human immune/haema
21	154.5	14.9	81	21	AAB42965	Human OREF 2729
22	150	14.5	377	22	AAM42051	Human polypeptide
23	145	14.0	374	22	AAM40959	Human polypeptide
24	145	14.0	374	22	AAM40960	Human polypeptide
25	144.5	13.9	360	22	ABB66833	Drosophila melanog
26	144.5	13.9	399	22	ABB61169	Drosophila melanog
27	143.5	13.9	449	22	ABB59263	Drosophila melanog
28	142	13.7	413	22	ABB69108	Drosophila melanog
29	138	13.3	329	22	AAM42411	Human polypeptide
30	137	13.3	329	22	AAM42413	Human polypeptide
31	137	13.2	125	22	AAU31107	Novel human secret
32	136.5	13.2	304	21	AAB50384	Human uncoupling p
33	136	13.1	305	22	AAM40265	Human polypeptide
34	135.5	13.1	285	21	AAG16697	Arabidopsis thalia
35	135.5	13.1	320	21	AAB50379	Human uncoupling p
36	135.5	13.1	320	22	AAM79071	Human protein sequ
37	135.5	13.1	320	22	AAM93892	Human polypeptide,
38	135.5	13.1	335	21	AAB51795	Human secreted pro
39	135.5	13.1	363	21	AAG27856	Arabidopsis thalia
40	131.5	12.7	297	22	ABB60934	Drosophila melanog
41	131.5	12.7	317	22	ABB62682	Drosophila melanog
42	131.5	12.7	320	22	AAM39031	Human polypeptide
43	131	12.6	310	22	AAB93001	Human protein sequ
44	131	12.6	311	21	AAB50382	Human uncoupling p
45	131	12.6	311	22	AAM39173	Human polypeptide

ALIGNMENTS

RESULT 1
AAM41505
ID AAM41505 standard; Protein; 318 AA.
XX
AC AAM41505;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6436.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW leukoemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0489725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX

XX PA (INCYTE GENOMICS INC.
 XX PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 XX PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
 XX PI WPI: 2001-041424/05.
 DR N-PSDB; AAF27733.
 XX PT Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX Claim 2; Page 133-134; 165pp; English.
 XX CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTS). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 XX SQ Sequence 289 AA;

Query Match 60.2%; Score 624; DB 22; Length 289;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-54;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 MVAGAVAGILEHCVMYPIDCVKTRMSLOPPDPAARYRNVLALWRIIRTEGLWRPMLN 135
 DB 1 mvagavagilehcvmypidcvktrmgsldpdpaaaryrnvlalwriirteglwrpmln 60
 QY 136 VTATGAGPAHALYFACYEKIKKTLSDVTHPGGNSHANGAGCVATLLHDAAMPAAE 192
 DB 61 vtatgagpahalyfacyekikktlsdvthpggnsiangagcgvatllhdaampae 117

RESULT 4
 AAM79039
 ID AAM79039 standard; Protein; 299 AA.
 AC AAM79039;
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1701.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 XX WO200157190-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US04098.
 XX 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 XX 20-JUN-2000; 2000US-0598075.
 XX 19-JUL-2000; 2000US-0620325.
 XX 01-SEP-2000; 2000US-0654936.
 XX 15-SEP-2000; 2000US-0663561.
 XX 20-OCT-2000; 2000US-0693325.
 XX 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK52172.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 20; Page 4043; 622lpp; English.
 XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX SQ Sequence 299 AA;

Query Match 49.9%; Score 516.5; DB 22; Length 299;
 Best Local Similarity 55.7%; Pred. NO. 1.3e-43;
 Matches 107; Conservative 19; Mismatches 39; Indels 27; Gaps 4;
 QY 1 MELRGAGGVAGGPAAGRSPGESALLDGLQGVGRGAGGAGACRPVQDPDPSG 60
 DB 28 melr---sgsv--gsqavarrmdgds-----rdgggg-----kdatgs 60
 QY 61 PDYEALPAGATVTTTHMVAGAVAGILEHCVMYPIDCVKTRMSLOPPDPAARYRNVLALWR 120
 DB 61 edyenlptsasvsthtmagamagilehsvmypdvsvktrmgsldpdpkaqytsvvalgk 120
 QY 121 IIRTEGLWRPMLNVTATGAGPAHALYFACYEKIKKTLSDVTHPGGNSHANGAGCVA 180
 DB 121 imrtgegrwprlrgvnmimgagpahamyfacyenmkrtlndvfhqgnshtangagsma 180
 QY 181 TLLHDAAMPAAE 192
 DB 181 tllhdavmpae 192

RESULT 5
 AAB60658
 ID AAB60658 standard; Protein; 331 AA.
 AC AAB60658;
 DT 04-MAY-2001 (first entry)
 DE Human mitochondrial solute carrier protein hMSC-o.
 KW Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
 KW preparation; detection.
 XX Homo sapiens.
 XX CNL269409-A.
 XX 11-OCT-2000.
 XX 17-MAR-2000; 2000CN-0114958.
 XX 17-MAR-2000; 2000CN-0114958.
 XX (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX Zhang X, Gao X, Xiao H;
 PI WPI: 2001-050544/07.
 DR N-PSDB; AAF59920.
 XX
 PT New human mitochondrion solute carrier protein and its nucleic acid -
 PS Claim 4; Page 20; 21pp; Chinese.
 XX
 CC The invention relates to a novel human mitochondrial solute carrier
 CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
 CC expressed in normal human hypothalamus tissue. The invention also relates
 CC to the preparation of hMSC-o proteins and nucleic acids, and the
 CC detection of hMSC-o proteins and nucleic acids in a sample. The present
 CC sequence represents hMSC-o.
 XX
 SQ Sequence 331 AA;

Query Match 49.7%; Score 515; DB 22; Length 331;
 Best Local Similarity 61.8%; Pred. No. 2.1e-43;
 Matches 97; Conservative 18; Mismatches 32; Indels 10; Gaps 1;

QY 36 GVGGRAGGAGACRPVPRQDPDSCPDYEALPAGATVTTTHMVAGAVAGILEHCVMPIDC 95
 Db 3 gdsrdggg-----kdatgsedyenlptsasvsthtmagamagilehsvmpvds 52
 QY 96 VKTRMQSLQDPDAARYRNVLALRIIRTEGLWRPMPGLNVTATGAGPAHALYFACYEKL 155
 Db 53 vktrmqslspdpkagqtslygalkkrmtegtwprlgvnmimgagpahamyfacyennm 112
 QY 156 KKTSLDVHPGNSHIANGAAGCVATLLHDAAMNPAE 192
 Db 113 krtldvfhqgnshtlangiagsmatllhdavmnpae 149

RESULT 6
 AAM39748
 ID AAM39748 standard; Protein; 677 AA.
 XX
 AC AAM39748;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #239.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-611725/70.
 XX
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 PS Claim 20; Page 189; 765pp; English.
 XX

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAM29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 677 AA;

Query Match 47.0%; Score 486.5; DB 22; Length 677;
 Best Local Similarity 53.1%; Pred. No. 3.6e-40;
 Matches 102; Conservative 20; Mismatches 43; Indels 27; Gaps 4;

QY 1 MELEGRGAGGVAGGPAAGPGRSPGESALLDGLQGVGRGAGGEGACRPVPRQDPDSCG 60
 Db 13 melr---sgsv--gsqavarmdgd-----rdgggg-----kdatgs 45
 QY 61 PDEALPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPDAARYRNVLALWR 120
 Db 46 edyenlptsasvsthtmagamagilehsvmpvdsvktrmqslspssqspvsiygalkk 105
 QY 121 IIRTEGLWRPMPGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGNSHIANGAGCVA 180
 Db 106 lmrtegtwprlgvnmimgagpahamyfacyennmkrtdvfhqgnshtlangiagsma 165
 QY 181 TLLHDAAMNPAE 192
 Db 166 tlhdavmnpae 177

RESULT 7
 AAM39719
 ID AAM39719 standard; Protein; 268 AA.
 XX
 AC AAM39719;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2864.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AA158875.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Example 4; SEQ ID NO 2864; 10078pp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 268 AA;

Query Match 46.0%; Score 477; DB 22; Length 268;
Best Local Similarity 94.7%; Pred. No. 1e-39;
Matches 90; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 TRMQSLQPPAARYNVLEALWRIIRTEGLWRPMDGLNVTATGAGPAHALYFACYEKLKK 157
Db 2 trmhsqspagrynvlevlwrirrtglwrpmdglntatgawpahalyfacyekllk 61

QY 158 TLDSDVHPGNSHIANGAGCVATLLHDAAMNPAE 192
Db 62 tldsdvhpqnshtangaagcvtatllhdaamnpae 96

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX
AC AAM00938;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow protein, SEQ ID NO: 414.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX
PN WO200153453-A2.
XX
PD 26-JUL-2001.
XX
PF 23-DEC-2000; 2000WO-US34960.

XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
DR N-PSDB; AAH90057.
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX Claim 10; Page 504-505; 648pp; English.
XX The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 366 AA;

Query Match 44.9%; Score 465; DB 22; Length 366;
Best Local Similarity 70.0%; Pred. No. 2.4e-38;
Matches 84; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

QY 73 TTHVAGAVAGILHRCVMTPIDCVKTRMSLQPPAARYNVLEALWRIIRTEGLWRPMDR 132
Db 74 sthmtagmagilehsvmypdvsktrmqslspdpkagysiygalkkimrtgfwrpir 133

QY 133 GLNVTATGAGPAHALYFACYEKLKTLSDVHPGNSHIANGAGCVATLLHDAAMNPAE 192
Db 134 gvnvmimggppahamyfacyenmkrtindvfhqgnshtangagsmatllhdavmnpae 193

RESULT 9
AAB42966
ID AAB42966 standard; Protein; 155 AA.
XX
AC AAB42966;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX OR2730 polypeptide sequence SEQ ID NO:5460.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipariatic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 272 AA;

Query Match 35.0%; Score 363; DB 21; Length 272;

Best Local Similarity 67.0%; Pred. No. 2.9e-28;

Matches 65; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 96 VKTRMSQLQDPDAARYNVLEALWRIETGLWRPRLNVTATGAGPAHALYFACYEKL 155

DB 3 lgrmslspdkqysiygalkmrtegrfwrgrvnmimgagpahamfacyenm 62

QY 156 KKTLSVDVHPGNSHANGAAGCVATLLHDAAMNPAE 192

DB 63 krtldvfhqgshiangiagsgmatllhdavmpae 99

RESULT 11

ABB61130

ID ABB61130 standard; Protein; 379 AA.

XX AC ABB61130;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 10182.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL05233.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure; SEQ ID NO 10182; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 379 AA;

Query Match 28.5%; Score 295; DB 22; Length 379;

Best Local Similarity 47.3%; Pred. No. 2.9e-21;

Matches 62; Conservative 22; Mismatches 41; Indels 6; Gaps 3;

QY 62 DYEALPAGATVTTMVGAVAGILEHCVMYPIDCVKTRMSQLQDPDAARYNVLEALWRI 121

DB 5 dyeslpt-tsvgnmtagaiaagvlehvmypldsvktrmsql--sptkmmnvlstlrmt 61

QY 122 IRTEGLWRPRLNVTATGAGPAHALYFACYEKLKKTLSVDVHPGNSHANGAAGCVAT 181

DB 62 itregllrpirgasavvlgagpahslyfaayemtkeltakftsvrnlnnyisga---vat 118

QY 182 LLHDAAMNPAE 192

DB 119 lhdaispsptd 129

RESULT 12

ABB51123

ID ABB51123 standard; Protein; 126 AA.

XX AC ABB51123;

DT 07-FEB-2002 (first entry)

XX Human secreted protein encoded by gene 163 SEQ ID NO:1076.

DE Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;

XX dermatological; immunosuppressive; antiinflammatory; immunostimulant;

KW cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;

KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnery;

KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;

KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;

KW human immunodeficiency virus; hyperproliferative disorder; wound healing;

KW Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;

KW Chaga's cardiomyopathy; coronary arteriosclerosis; angioenic disorder;

KW corneal graft neovascularisation; diabetic retinopathy; regeneration;

KW neurological disorder; Huntington's chorea; Alzheimer's disease;

KW Parkinson's disease; infectious disease; chromosome 8.

XX Homo sapiens.

XX WO200162891-A2.

XX 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US05614.

XX 24-FEB-2000; 2000US-184836P.

PR 29-MAR-2000; 2000US-193170P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;

PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;

PI Florence C, Hu J, Kyaw H, Fischer CL, Ferrie AM, Fan P;

PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;

PI Zeng Z, Greene JM;

XX WPI; 2001-625724/72.

DR Nucleic acids encoding 207 human secreted polypeptides, useful for

XX preventing, diagnosing and/or treating, e.g. cancers, Parkinson's

PT disease and diabetic retinopathy -

PT Disclosure; Page 357; 1533pp; English.

XX ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted

CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various

CC activities based on the tissues and cells the genes are expressed in.

CC Example of these activities include: immunomodulatory; antisclerotic;

CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;

CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;

CC neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;

CC antiparkinsonian; antimicrobial; and vulnery. (I) and (II) can be used

CC in gene therapy and vaccine production. (I) and (II) can be used in the

CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA8185 to
CC ABA83193 and ABB50300 represent sequences used in the exemplification of
CC the present invention.

XX Sequence 126 AA;

Query Match 23.3%; Score 241.5; DB 22; Length 126;
Best Local Similarity 48.6%; Pred. No. 1.8e-16;
Matches 54; Conservative 11; Mismatches 23; Indels 23; Gaps 3;

Qy 36 GVGRCAGGGEAGACRPPVRQDPDSQPDYALPAGATVTHMVAGAVAGILEHCVMYPIDC 95

Db 36 gdsrdgggg-----katsgedyenlptsasvsthtagmagilehsvmypyds 85

Qy 96 VKTRMQSLQPPAARYRNVLALMRIITEGLWRPMGLNVTATGAGPAHA 146

Db 86 vktrmqslspdpkadytsiygalkkkmrte-----asg-gpcea 123

RESULT 13

AAG22077

ID AAG22077 standard; Protein: 331 AA.

XX AC AAG22077;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24869.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX XX 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

```
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.5%; Score 223; DB 21; Length 331;
Best Local Similarity 39.9%; Pred. No. 4e-14;
Matches 59; Conservative 22; Mismatches 51; Indels 16; Gaps 5;

QY 52 PVRODPDGDYEALPAGATVTT-----HMVAGAVAGILEHCVMYPIDCVKTRMQSLOPD 106
Db 16 pipppdfhp---aivpaqnttlkfqlmvgagsvehmanfpvdtvktmqlrsc 72
QY 107 PAARYNVLEALWRIIRTEGLWRPMRLNVTATGAGPAHALYFACYEKLKLTLSVDVHPG 166
Db 73 p-ikpigrqfarslktktdpsalrygiwanglgagbahavfyfsvevskfls-----g 126
QY 167 G--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 127 gnpnsaahaisgvfatissdavftpmd 154

RESULT 14
AAG43092
ID AAG43092 standard; Protein; 331 AA.
XX
AC AAG43092;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53820.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
```

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0140353.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 21.5%; Score 223; DB 21; Length 331;
Best Local Similarity 39.9%; Pred. No. 4e-14;
Matches 59; Conservative 22; Mismatches 51; Indels 16; Gaps 5;
QY 52 PVRQDPDSDPYEALPAGATVTT-----HMVAGAVAGILEHCVMYPIDCVKTRMQSLQPD 106
Db 16 pibqpdpdhp---aaiivpaqnttlkfwlmvagsiagsvehmamfpvdtvktmqlrsc 72
QY 107 PAARYRNVLALWRIIRTEGLWRPMBGLNVTATGAGPAHALYFACYEKLKLTSLSDVIHPG 166
Db 73 p-ikpigrqafsrilktgpsalyrgiawmglgagpahavfysfeyvskkfls-----g 126
QY 167 G--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 127 gnpnnsaahaisgvfatissdavftpm 154
RESULT 15
AAW17054
ID AAW17054 standard; Protein; 289 AA.
AC AAW17054;
XX
XX 05-AUG-1997 (first entry)
XX Blackcurrant RIB7 polypeptide.
XX Blackcurrant; fruit-specific promoter; RIB7; transgenic plant.
XX Ribes nigrum strain Ben Alder.
XX WO9717452-Al.
XX 15-MAY-1997.
XX 04-NOV-1996; 96WO-EP04807.
XX 03-NOV-1995; 95GB-0022558.
XX (SMTK) SMITHKLINE BEECHAM PLC.
XX Brennan RM, Taylor MA, Woodhead MR;
PI
XX
XX WPI: 1997-281041/25.
XX N-PSDB; AAT68957.
XX New isolated promoters from blackcurrant fruit - used for driving
PT fruit-specific expression of DNA sequences in transgenic
PT blackcurrant and other non-climacteric fruit
XX
XX Example 4; Page 35-36; 66pp; English.
XX
XX RIB polypeptides (AAW17050-54) are encoded by cDNA clones (AAT68953-
CC 57) that exhibit differential expression in blackcurrant fruit
CC during the ripening period of fruit development. RIB7 shows
CC 62% similarity to yeast MRS4, a yeast mitochondrial RNA splicing
CC protein. RIB7 is expressed almost entirely in fruit. The
CC promoter region (AAT68952) of the RIB7 gene can be used as a
XX fruit-specific promoter.
XX

SQ Sequence 289 AA;
Query Match 20.4%; Score 211; DB 18; Length 289;
Best Local Similarity 38.5%; Pred. No. 5.5e-13;
Matches 45; Conservative 26; Mismatches 42; Indels 4; Gaps 2;
QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVLALWRIIRTEGLWRPMBGLN 135
Db 1 miagsiagsiehmamypvdtlkttriqai-gscsaqsaglrqalgsilkvegpaglyrgig 59
QY 136 VTATGAGPAHALYFACYEKLKLTSLSDVIHPGNSHIANGAAGCVATLLHDAAMNPAE 192
Db 60 amglgagpahavfysfeyemcketfs---hgdpsnsgahavsgvfatvasdavitpmd 113
Search completed: August 27, 2002, 03:55:58
Job time: 4910 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:49 ; Search time 48.11 Seconds
(without alignments)
97.987 Million cell updates/sec

Title: US-09-870-113-4

Perfect score: 1036

Sequence: 1 MELEGRGAGGVAGGPAAGPG.....GAAGCVATILLHDAMNPAEG 193

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	212	20.5	328	4	US-09-068-140A-15
2	211	20.4	289	4	US-09-068-140A-10
3	135.5	13.1	320	2	US-08-933-750C-12
4	135.5	13.1	320	4	US-09-234-613-12
5	131	12.6	351	2	US-08-933-750C-19
6	131	12.6	351	4	US-09-234-613-19
7	124.5	12.0	311	2	US-08-775-009-32
8	120	11.6	311	2	US-08-775-009-33
9	115	11.1	447	4	US-09-160-119-4
10	115	11.1	674	4	US-09-160-119-2
11	112.5	10.9	306	5	PCT-US94-09799-1
12	106.5	10.3	469	4	US-09-188-930-339
13	102.5	9.9	432	2	US-08-937-466-4
14	102.5	9.9	432	2	US-09-172-528-4
15	102.5	9.9	432	3	US-09-318-199-4
16	102.5	9.9	432	4	US-09-503-579-4
17	98.5	9.5	308	2	US-08-937-466-2
18	98.5	9.5	308	2	US-09-172-528-2
19	98.5	9.5	308	3	US-09-318-199-2
20	98.5	9.5	308	4	US-09-503-579-2
21	97	9.4	180	4	US-08-937-263B-8
22	94.5	9.1	312	4	US-09-142-565-2
23	94	9.1	180	2	US-08-791-495-9
24	94	9.1	312	4	US-09-188-930-142
25	93.5	9.0	256	2	US-08-937-466-6
26	93.5	9.0	256	2	US-09-172-528-6
27	93.5	9.0	256	3	US-09-318-199-6

Sequence 6, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 105, App

28 93.5 9.0 256 4 US-09-503-579-6
29 92.5 8.9 307 2 US-08-807-861A-56
30 92.5 8.9 307 3 US-09-210-681-56
31 92.5 8.9 307 3 US-08-946-719A-56
32 92.5 8.9 309 1 US-08-518-878B-51
33 92.5 8.9 309 2 US-08-807-861A-51
34 92.5 8.9 309 3 US-08-470-868A-51
35 92.5 8.9 309 3 US-09-210-681-51
36 92.5 8.9 309 3 US-08-946-719A-51
37 91 8.8 298 3 US-08-961-871-10
38 89.5 8.6 103 1 US-08-209-747-4
39 89.5 8.6 103 1 US-08-458-298-4
40 89.5 8.6 299 1 US-08-518-878B-56
41 89.5 8.6 299 2 US-08-470-868A-56
42 88.5 8.5 1332 1 US-07-609-716-41
43 88.5 8.5 1332 4 US-08-475-411A-41
44 88.5 8.5 1332 4 US-08-478-029A-41
45 87.5 8.4 146 1 US-07-609-716-105

ALIGNMENTS

RESULT 1
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-15

QY 138 ATGAGPAHALYFACYEKLKKTLL-----SDVIHPGNSHIANGAGCVATLLHDAAMNPA 191
||:|:| |:|:| | ||: | ||: | :|
Db 123 LMTVPATAIYFYAYDQLKAFGLCGRALTSPLYAP-----MVAGALARLGTVTVISPL 174
||:|:| |:|:| | ||: | ||: | :|
QY 192 E 192
|
Db 175 E 175
|
RESULT 6
US-09-234-613-19
; Sequence 19, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNOCAT01
; CLONE: 724157
US-09-234-613-19
Query Match 12.6%; Score 131; DB 4; Length 351;
Best Local Similarity 24.9%; Pred. No. 9.1e-06;
Matches 45; Conservative 19; Mismatches 51; Indels 66; Gaps 5;
QY 67 PAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPA-----109
||: | ||: | :|
Db 6 PAGISPLQMVASGTGAVVTSLFMTPLDVVKVLRQSRPMSASELMPSSRLWSLSYTKWK 65
||: | ||: | :|
QY 110 -----RYRNVLEALWRIITEG---LWRPMTGLNVT 137
|: :|:| |:| | || |

Db 66 CLLYCNGVLEPLYLCPNGARCATWFOQDPRFTCTMDAFVKIVRHEGTRTLM---SGLPAT 122
QY 138 ATGAGPAHALYFACYEKLKKTLL-----SDVIHPGNSHIANGAGCVATLLHDAAMNPA 191
||:|:| |:|:| | ||: | ||: | :|
Db 123 LMTVPATAIYFYAYDQLKAFGLCGRALTSPLYAP-----MVAGALARLGTVTVISPL 174
||:|:| |:~:~:~: | ||: | ||: | :|
QY 192 E 192
|
Db 175 E 175
|
RESULT 7
US-08-775-009-32
; Sequence 32, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Wellong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Velocardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-775-009-32
Query Match 12.0%; Score 124.5; DB 2; Length 311;
Best Local Similarity 31.4%; Pred. No. 3.7e-05;
Matches 32; Conservative 22; Mismatches 43; Indels 5; Gaps 2;
QY 65 ALPA-GATVTTTH-----MVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYRNVLALW 119
||: | ||: | :|
Db 13 AAPASCKAKLTHPEKAILAGGLAGGIEICITFPTEYVKTQLQDLSHPPRYRGIGDCVR 72
||: | ||: | :|
QY 120 RIIRTGLWRPMTGLNVTATGAGPAHALYFACYEKLKKTLLSD 161
||: | ||: | :|
Db 73 QTVRSHGVGLYRGLSLLYGSIPKAAVRFGMFEFLSNHMRD 114
||: | ||: | :|
RESULT 8
US-08-775-009-33
; Sequence 33, Application US/08775009

; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weillong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Velicardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775.009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-775-009-33

Query Match 11.6%; Score 120; DB 2; Length 311;
Best Local Similarity 29.6%; Pred. No. 0.00011;
Matches 29; Conservative 23; Mismatches 42; Indels 4; Gaps 1;
QY 68 AGATVTH----MWAGVAGILEHCVMPIDCVKTRMOSLQDPDPAARYNVLEALWRIIR 123
DB 17 SGKALTHPGKAILAGGLAGGTEICITFTYVKTQLQDLDERANPPRYRGIGDCVQTVR 76
QY 124 TEGLRWPRMGLNVTATGAGPAHALYFACYEKLKLTSLD 161
DB 77 SHGVGLYRGUSSLLYGSIPKAAVRFGMFELSNHRD 114
RESULT 9
US-09-160-119-4
; Sequence 4, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRILL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160.119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 447
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-4

Query Match 11.1%; Score 115; DB 4; Length 447;
Best Local Similarity 30.1%; Pred. No. 0.00057;
Matches 34; Conservative 18; Mismatches 39; Indels 22; Gaps 4;
QY 79 GAVAGILEHCVMPIDCVKTRMOSLQDPDPA----ARYRVNLEALWRIIRTEGLWRPMRGL 134
DB 107 GSVAGAVGATAVPIDLKTRMQNQRSTGSGFVGMELMYKNSFDCFKVLRYEGFFGLYRGL 166
QY 135 NVTATGAGPAHALYFACYEKLKLTSLDVI-----HPGGN-----SHIANGAAG 177
DB 167 LPQLLGVAPEKA-----IKLTVNDVYRDKFMHKDGSVPLAAETLAGGCAG 211

RESULT 10
US-09-160-119-2
; Sequence 2, Application US/091601119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRILL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160.119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-2

Query Match 11.1%; Score 115; DB 4; Length 674;
Best Local Similarity 30.1%; Pred. No. 0.00098;
Matches 34; Conservative 18; Mismatches 39; Indels 22; Gaps 4;
QY 79 GAVAGILEHCVMPIDCVKTRMOSLQDPDPA----ARYRVNLEALWRIIRTEGLWRPMRGL 134
DB 334 GSVAGAVGATAVPIDLKTRMQNQRSTGSGFVGMELMYKNSFDCFKVLRYEGFFGLYRGL 393
QY 135 NVTATGAGPAHALYFACYEKLKLTSLDVI-----HPGGN-----SHIANGAAG 177
DB 394 LPQLLGVAPEKA-----IKLTVNDVYRDKFMHKDGSVPLAAETLAGGCAG 438

RESULT 11
PCT-US94-09799-1
; Sequence 1, Application PC/TUS9409799
; GENERAL INFORMATION:
; APPLICANT: Kagan, David
; TITLE OF INVENTION: Method and Composition for Weight Reduction
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Watov & Kipnes, P.C.
; STREET: 186 Princeton-Hightstown Rd, PO Box 247
; CITY: Princeton Junction
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08550

```

1  GENERAL INFORMATION.
2  APPLICANT: Watson, James D.
3  APPLICANT: Strachan, Lorna
4  APPLICANT: Sleeman, Matthew
5  APPLICANT: Onrust, Rene
6  APPLICANT: Murison, James Greg
7  TITLE OF INVENTION: Compositions Isolated From Skin Cells
8  TITLE OF INVENTION: and Methods For Their Use
9  FILE REFERENCE: 11000.1011c1
10 CURRENT APPLICATION NUMBER: US/09/188,930A
11 CURRENT FILING DATE: 1998-11-09
12 NUMBER OF SEQ ID NOS: 348
13 SOFTWARE: FastSEQ for Windows Version 3.0
14 SEQ ID NO 339
15     LENGTH: 469
16     TYPE: PRT
17     ORGANISM: Mouse
18 US-09-188-930-339

```

RESULT 14
US-09-172-528-4


```

RESULT 15
US-09-318-199-4
; Sequence 4, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA

```

Search completed: August 27, 2002, 03:58:50
Job time: 4892 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:57:40 ; Search time 80.04 Seconds
(without alignments)
231.700 Million cell updates/sec

Title: US-09-870-113-4

Perfect score: 1036

Sequence: 1 MELEGRGAGVGAGPRAAGP.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	338.5	32.7	312	2 T26089	hypothetical prote
2	252	24.3	303	2 T39149	probable RNA splic
3	251.5	24.3	314	2 S5179	mRNA splice defec
4	224.5	21.7	781	2 A86205	hypothetical prote
5	223.5	21.6	304	2 S13533	mRNA splice defec
6	223	21.5	331	2 T00582	probable mitochond
7	168	16.2	307	2 S00949	probable phosphate
8	150	14.5	310	2 T20229	hypothetical prote
9	143	13.8	358	2 T45934	hypothetical prote
10	142	13.7	348	2 D84798	probable mitochond
11	134.5	13.0	436	2 J01459	Btl protein precu
12	132.5	12.8	291	2 T37992	probable tricarbox
13	131	12.6	351	2 T34393	hypothetical prote
14	129	12.5	313	2 D84613	hypothetical prote
15	128.5	12.4	332	2 T47703	Ca-dependent solu
16	128	12.4	352	2 T01729	mitochondrial solu
17	124.5	12.0	296	2 B96830	hypothetical prote
18	124.5	12.0	311	2 G01789	citrate transporte
19	124	12.0	301	1 S31935	ADP,ATP carrier pr
20	124	12.0	330	2 T09362	hypothetical prote
21	124	12.0	902	1 S54495	probable carrier p
22	123	11.9	367	2 G96770	hypothetical prote
23	122	11.8	392	2 T05350	adenylate transloc
24	121.5	11.7	310	2 S69050	probable membrane
25	121.5	11.7	313	2 T05577	uncoupling protein
26	121	11.7	650	2 T32897	hypothetical prote
27	120.5	11.6	322	2 S57116	probable carrier p
28	120	11.6	311	2 A46595	tricarboxylate tra
29	119.5	11.5	311	2 G86383	probable mitochond

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26089

R:Swineburne, J.; Ainscough, R.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z20149

A:Accession: T26089

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-312 <WIL>

A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9

A:Experimental source: clone W02B12

C:Genetics:

A:Gene: CESP:W02B12.9

A:Map position: 2

A:Introns: 18/3; 251/3; 286/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 32.7%; Score 338.5; DB 2; Length 312;

Best local Similarity 48.0%; Pred. No. 8.8e-21;

Matches 73; Conservative 23; Mismatches 37; Indels 19; Gaps 4;

QY 42 GGGEAGACRPVVRQDPDSDGPDYEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQ 101

Db 3 GGGE-----DEYESLPT-HSVPVHLTAGALAGAVEHCVMFPDPSVKTRMQ 46

QY 102 SIQPPDPAARYNVLEALWRIIRTEGLWRPMRGLNTATGAGPAHALYFACYEKLKLTSD 161

Db 47 SLCPCPETKCPVHSLMSIVKREGWLRLRGVNAAGSMPAHALYFTVYERKMGLTG 106

QY 162 VIHPGNSH-IANGAGCVATLLHDAAMNPAE 192

Db 107 --NSAGHSNTLAYAGSVVATLIHDAIMNPAE 136

RESULT 2

T39149

probable RNA splicing protein mitochondrial carrier protein - fission yeast (Schizosacc

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T39149

R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z21748

A:Accession: T39149

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-303 <OLI>

A;Cross-references: EMBL:Z99168; PIDN:CAB1300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A;Experimental source: strain 972h-; cosmid c8C9
C;Genetics:
A;Gene: SPDB:SPAC8C9.12c
A;Map position: 1
C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 24.3%; Score 252; DB 2; Length 303;
Best Local Similarity 40.3%; Pred. No. 1.3e-13;
Matches 52; Conservative 25; Mismatches 48; Indels 4; Gaps 1;

Qy 62 DYELPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMOSLQDPDPAARYNLEALWRI 121
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 9 DYELGPIGSPYAHLLAGAFSGILEHSVMYPVDAIKTRMQLNGVSRVSGNIYNSVIKI 68

Qy 122 IRTGLWRPRLNVTATGACPAHALYFACYEKLKLTLSVDVHPGGNSHTANGAGCVAT 181
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 69 SSTEGVYSLWRGISSVINGAGPSHAIFYVLEFFKSK----INASPDRLASALAGACAI 124

Qy 182 LLHDAAMNP 190
: ||| |
Db 125 TISDAFWTP 133

RESULT 3
S55179
mRNA splice defect-suppressing mitochondrial carrier MRS3 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J0675; protein YJL133w
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence.revision 08-Sep-1995 #text_change 21-Jul-2000
C;Accession: S55179; S01267; S20228; S56915; S71664
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995
A;Description: The complete sequence of a 40.7 kb segment located on the left arm of yeast hypothetical proteins.
A;Reference number: S55159
A;Accession: S55179
A;Molecule type: DNA
A;Residues: 1-314 <KAT>
A;Cross-references: EMBL:X87371; NID:g854542; PID:g854563
R;Schmidt, C.; Soellner, T.; Schweyen, R.J.
Mol. Gen. Genet. 210, 145-152, 1997
A;Title: Nuclear suppression of a mitochondrial RNA splice defect: nucleotide sequence and reference number: S01267; MUID:88121698
A;Accession: S01267
A;Molecule type: DNA
A;Residues: 1-86, 'S', 88-102, 'L', 104-127, 'Y', 129-141, 'M', 143-187, 'R', 189-314 <SCH>
A;Cross-references: EMBL:X06239; NID:g3990; PIDN:CAA29582.1; PID:g3991
R;Wiesenberger, G.; Link, T.A.; von Ahsen, U.; Waldherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A;Title: MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the protein sequence database
A;Reference number: S13532; MUID:91108815
A;Accession: S20228
A;Molecule type: DNA
A;Residues: 13-314 <WIE>
A;Cross-references: EMBL:X56445; NID:g3992; PIDN:CAA39830.1; PID:g3994
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56912
A;Accession: S56915
A;Molecule type: DNA
A;Residues: 1-314 <KAW>
A;Cross-references: EMBL:Z49408; NID:g1008337; PID:g1008338; MIPS:YJL133w
R;Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X of chromosome XI.
A;Reference number: S71643; MUID:96408771
A;Accession: S71664
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-314 <KAF>

A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CRAA0822.1; PID:g854563
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: SGD:MRS3
A;Cross-references: SGD:S0003669; MIPS:YJL133w
A;Map position: 10L
A;Genome: nuclear
C;Function:
A;Description: probably involved in splicing of all intron for COB gene; essential for licing
A;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C;Keywords: mitochondrial inner membrane; mitochondrion; pre-mRNA splicing
F;30-119/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F;127-211/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F;216-311/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 24.3%; Score 251.5; DB 2; Length 314;
Best Local Similarity 41.0%; Pred. No. 1.4e-13;
Matches 57; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

Qy 52 PVRODPDGPDYELPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMOSLQDPDPAARY 111
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 13 PIPAIPMDLPDYELPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMOSLQDPDPAARY 71

Qy 112 RNVEALWRIIRTEGLWRPRLNVTATGACPAHALYFACYEKLKLTLSVDVHPGGNSHI 171
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 72 KNMLSQISHISTSEGLALNKGVSQVILGAGPAHAVFGYEFCKKILIDSSDQTTHHPF 131

Qy 172 ANGAAGCVAILLHDAAMNP 190
: ||| |
Db 132 KTAISGACATTASDALMNP 150

RESULT 4
A86205
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence.revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A86205
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.K.; Creasy, T.H.; Dewar, C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: A86205
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-781 <STO>
A;Cross-references: GB:AE005172; NID:g8954043; PIDN:AAF82217.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 21.7%; Score 224.5; DB 2; Length 781;
Best Local Similarity 38.0%; Pred. No. 6.4e-11;
Matches 57; Conservative 23; Mismatches 55; Indels 15; Gaps 5;

Qy 52 PVRODPD-----SGPDY--BALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMOSLQ 104
| : ||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 9 PKTEPDLQVQSOTDPDKPEIADHGLKFWQFMTAGSAGSVEHWMAMPVDTIKTHMOALR 68

Qy 105 PDPAARYRNVEALWRIIRTEGLWRPRLNVTATGACPAHALYFACYEKLKLTLSVDVH 164
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 69 PCP-LKPVGIREAFRSIIQEGPSALYRGIWAMGLGACPAHAVYFSFYESSKYL----- 123

Qy 165 PGG--NSHIANGAAGCVATLLHDAAMNPAE 192

A:Cross-references: EMBL:X92441; NID:g1050762; PIDN:CAA63185.1; PID:g1050774
C:Genetics:
A:Map position: 15R
A:Note: YOR222w
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:9-107/Domain: ADP,ATP carrier protein repeat homology <ACPI>
F:115-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:208-300/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 16.2% Score 168; DB 2; Length 307;
Best Local Similarity 33.9%; Pred. No. 1.2e-06;
Matches 42; Conservative 20; Mismatches 54; Indels 8; Gaps 1;

QY 77 VAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAA-----RYRNVLLEALWRIIRTEGLW 128
 :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 17 ISGAVAGISELTVMYPLDVVKTRFQLEVTFPTAAAGVQVERYNVIDCLKKIYKKEGFS 76

QY 129 RPMGLNVATGAGPAHALFYACEKKUKTSDVIHPGGNSHIANGAACGVATLLHDDAM 188
 | ||| | | | | | ||| | : | : | : | : | : | : | : | : | : | : | : |
Db 77 RLYRGISSPMLMEAPKRATKFACDQYQIKFNLFNTNETTKISIAAGASAGMTAAVI 136

QY 189 NPAE 192
 | |
Db 137 VPFE 140

RESULT 8
T20229
hypothetical protein C54G10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20229
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20229
A>Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-310 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99811.1; GSPDB:GN00023; CESP:C54G10.4
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.4
A:Map position: 5
A:Introns: 14/1; 63/1; 155/2; 256/2; 299/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

R;Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23017
A:Accession: T45934
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <MON>
A:Cross-references: EMBL:AL132960
A:Experimental source: cultivar Columbia; BAC clone F5K20
C:Genetics:
A:Map position: 3
A:Introns: 98/3; 157/3; 207/3; 234/3; 258/2; 281/1
A>Note: F5K20.240
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 13.8%; Score 143; DB 2; Length 358;
Best Local Similarity 28.8%; Pred.No.0.00016;
Matches 36; Conservative 23; Mismatches 56; Indels 10; Gaps 2;

QY 68 AGATVTTTHVAGAVAGILEHCVMPIDCVKTRMOSLPDPAARYRNLYEALWRIIRTEGL 127
|| : || : || : || : || : || : || : || : || : || : || : || :
Db 172 AGVDISVHFVSGGLAGLTAAATVPDLVRLTSLA-----QGQHAFRTICREGI 222

QY 128 WRPMRGLNVTTATGAGPAHALFYACYEKLLKTLSDVIHPGNSHIANGAACGVATLLHDAA 187
:|| : || : || : || : || : || : || : || : || : || : || : || :
Db 223 LGLYKGLGATLLGVGPSLAISFAAYETTF-TFWLSHRPNDSNNAVSLGCGSLGIIVSSTA 281

QY 188 MNPAE 192
| :
Db 282 TFPLD 286

RESULT 10
D84798
probable mitochondrial carrier protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84798
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84798
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: GB:AE002093; NID:g4895195; PIDN:AAD32782.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g37890
A:Map position: 2
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: mitochondrion

Query Match	13.7%;	Score 142;	DB 2;	Length 348;
Best Local Similarity	32.8%;	Pred. No. 0.00019;		
Matches	40;	Conservative 20;	Mismatches 46;	Indels 16;
Gaps				
4;				
QY	75	HMVAGAVAGILEHCVMYPIDCVKTRMQSDPPAPARYRNVLEALNRIIRTEGLWPMRGL	134	
		: : : : : : : : : : : : : :		
Db	151	HFVSGGLAGTATAATATYPLDLVTRLAQR--NAIYVGIEHTFTICREEGILGYKGL	208	
		: : : : : : : : : : : : :		
QY	135	NVTATGAGPAHALYFACYEKLK-----KTLSDVIHPGGNSHIANGAGCVATL---LH	184	
		: : : : : : : : : : : : :		
Db	209	GATLLGVGPSLAINFAYESMKLFWHSHRPNSDLV---VSLVSGGLAGAVSSDDKLF	264	
		: : : : : : : : : : : : :		
QY	185	DA	186	
Db	265	DA	266	
RESULT	9			
T45934				
hypoetical protein F5K20.240 - Arabidopsis thaliana				
C:Species: Arabidopsis thaliana (mouse-ear cress)				
C:date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000				
C:Accession: T45934				

A: Gene: At2g22500
A: Map position: 2
C: Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology

Query Match	12.5%	Score 129;	DB 2;	Length 313;
Best Local Similarity	29.7%;	Pred. No. 0.002;		
Matches	41;	Conservative 31;	Mismatches 52;	Indels 14; Gaps 5;

QY	61	PDYEALPAGATVTHMVAGAVAGTILEHCVMYPCDVCTRMQS----	LQDDPAARVRLVLEA	117
		: : : :	: : : :	
Db	112	PETKTM-----LMKKIGAGAIGAAGVADPADVAMVQMAGRLPLTDRNYSKSLDA	167	
		: : : :	: : : :	
QY	118	LWRIIREGLWRPRLGNVATGAGPAHALVFCYKELKKTLSD--	VHPPGNSHI--ANG	174
		: : : :	: : : :	
Db	168	ITQWIREGVTSLWRGSSLIINRAMLVTSQLASYDSVKETILEKGLDKGLGTHVSASF	227	
		: : : :	: : : :	
QY	175	AAGCVATLLHDAAMNPAE	192	
		:	:	
Db	228	AAGFVASV-----ASNPDV	241	
		:	:	

RESULT 15

T47703

Ca-dependent solute carrier-like protein - Arabidopsis thaliana

N:Alternate names: protein Fli16.50

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Sep-2000

C:Accession: T47703

R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansgorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24473

A:Accession: T47703

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332 <BEN>

A:Cross-references: EMBL:AL161667

A:Experimental source: cultivar Columbia; BAC clone Fli16

C:Genetics:

A:Map position: 3

A:Introns: 63/3; 122/3; 172/3; 206/3; 230/2; 253/1

A:Note: Fli16.50

C:Superfamily: ADP.ATP carrier protein; ADP.ATP carrier protein repeat homology

Query Match	12.4%	Score 128.5;	DB 2;	Length 332;
Best Local Similarity	28.5%;	Pred. No. 0.0023;		
Matches	35;	Conservative 23;	Mismatches 52;	Indels 13; Gaps 4;
QY	75	HMVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDAARYRVN-LEALWRIIT----	EGLWR	129
		:	: :	:
Db	144	HFVAGGLAGITASATYPLDLVTRL-----AAQTKVIYSGIWHITLRSITTD	EGILG	196
		:	: :	:
QY	130	PMRGLNVATAGGAPAHALYFACYEKLIKTKTLDVTHPGGNSHIANGAAGCVATILLHDAAMN		189
		: : : :	: :	:
Db	197	LYRGLGTTLVGVPSTAFISFVIESLSRYWRST-RPHDSPIMVSLACGSLSGTASSTATF		255
QY	190	PAE	192	
		:		
Db	256	PLD	258	

Search completed: August 27, 2002, 03:57:41
Job time: 4947 sec


```

CC -----
DR EMBL; X56445; CAA39829.1; -.
DR EMBL; X56445; CAA39830.1; ALT_INIT.
DR EMBL; X06239; CAA29582.1; ALT_SEQ.
DR EMBL; X87371; CAA60822.1; -.
DR EMBL; Z49408; CAA89428.1; -.
DR PIR; S01267; S01267.
DR PIR; S20228; S20228.
DR SGD; S0003669; MRS3.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW mRNA processing.
FT REPEAT 19 120 BLOCK I (APPROXIMATE).
FT REPEAT 121 212 BLOCK II (APPROXIMATE).
FT REPEAT 213 314 BLOCK III (APPROXIMATE).
FT TRANSMEM 33 52 POTENTIAL.
FT TRANSMEM 93 112 POTENTIAL.
FT TRANSMEM 130 149 POTENTIAL.
FT TRANSMEM 185 204 POTENTIAL.
FT TRANSMEM 219 238 POTENTIAL.
FT TRANSMEM 285 298 POTENTIAL.
SQ SEQUENCE 314 AA; 34501 MW; 861145CB6E4EF321 CRC64;

Query Match 24.38; Score 251.5; DB 1; Length 314;
Best Local Similarity 41.08; Pred. No. 1.4e-13;
Matches 5; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

QY 52 PVRODPSGPDYEALPAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARY 111
   I : ||||| : : : ||||| : ||||| : ||||| : ||||| :
Db 13 PIPAIPMDLPDYEALPHAPLYHQLIAGAFAGIMEHSVMPFDALKTRIQSANAKLSA- 71

QY 112 RNVEALWRIITGLWRPMRGLNVTATGAGPAHALFYACYEKLKTKTLDVHPGNGNSHI 171
   I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 72 KNMLSQISHISTSEGTALMKGVQSVILGAGPAHAVYFGTYEFCCKNLIDSSDTQTHHPF 131

QY 172 ANGAAGCVATLLHDAANMP 190
   : || || || ||
Db 132 KTAISGACATTASDALNMP 150

RESULT 2
MRS4_YEAST
ID MRS4_YEAST STANDARD; PRT; 304 AA.
AC P23500;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial RNA splicing protein MRS4.
GN MRS4 OR YKR052C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1301;
RX MEDLINE=91108815; PubMed=1703236;
RA Wiesenberger G., Link T.A., von Ahsen U., Waldherr M., Schweyen R.J.;
RT "MRS4 and MRS3, two suppressors of mRNA splicing defects in yeast,
   are new members of the mitochondrial carrier family.";
RL J. Mol. Biol. 217:23-37(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Vissers S., Urrestarazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: MRS4 SUPPRESSES A MITOCHONDRIAL SPLICING DEFECT IN THE
CC FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING
CC ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN
CC THE MITOCHONDRION (POSSIBLY OF CATIONS). NOT ESSENTIAL.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

```

```

CC -----
CC inner membrane.
CC -|- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,
CC TO YEAST MRS3; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X56444; CAA39828.1; -.
DR EMBL; Z28277; CAA82130.1; -.
DR PIR; S13533; S13533.
DR SGD; S0001760; MRS4.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
KW mRNA processing.
FT REPEAT 9 110 BLOCK I (APPROXIMATE).
FT REPEAT 111 202 BLOCK II (APPROXIMATE).
FT REPEAT 203 304 BLOCK III (APPROXIMATE).
FT TRANSMEM 23 41 POTENTIAL.
FT TRANSMEM 83 102 POTENTIAL.
FT TRANSMEM 120 139 POTENTIAL.
FT TRANSMEM 175 194 POTENTIAL.
FT TRANSMEM 209 228 POTENTIAL.
FT TRANSMEM 275 288 POTENTIAL.
SQ SEQUENCE 304 AA; 33307 MW; 5ABBF985B547EDC CRC64;

Query Match 21.68; Score 223.5; DB 1; Length 304;
Best Local Similarity 38.88; Pred. No. 2.5e-11;
Matches 50; Conservative 24; Mismatches 54; Indels 1; Gaps 1;

QY 62 DYEALPAGATVTTTHVAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNVEALWRI 121
   I : ||||| : : : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 13 DYEALPSHAPLHSQLIAGAFAGIMEHSVMPFDALKTRVQAGLNKAAS-TGMISQISKI 71

QY 122 IRTEGLWRPMRGLNVTATGAGPAHALFYACYEKLKTKTLDVHPGNGNSHIANGAGCVAT 181
   I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 72 STMEGSMALMKGVQSVILGAGPAHAVYFGTYEFCCKARLISPEDMQTHQPMKLTALSGTIAT 131

QY 182 LLHDAANMP 190
   : || || ||
Db 132 IADALNMP 140

RESULT 3
ODC2_YEAST
ID ODC2_YEAST STANDARD; PRT; 307 AA.
AC Q99237;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update);>
DE Mitochondrial 2-oxodicarboxylate carrier 2.
GN ODC2 OR YOR222W OR YOR50-12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96437977; PubMed=8840505;
RA Galisson F., Dujon B.;
RT "Sequence and analysis of a 33 kb fragment from the right arm of
   chromosome XV of the yeast Saccharomyces cerevisiae.";
RL Yeast 12:877-885(1996).
RN [2]

```


Db 4 YDKPGDGNNTKFOVAVAGSVGLVTRALISPFVDVIXIRFQ-LQHERLSRSDPSAKYHGI 62
QY 115 LEALWRIIRTEG---LWRPMGLNVTATGAGPAHALYFACYEKLKTLSDVVIHPGG---- 167
Db 63 LQASRQILQEGPTAFWKGHVPAQILSIGYGAOVFLSF-----EMLTGLVHRGSVYDA 115

QY 168 ---NSH-IANGAGCVATL 182

Db 116 REFSVHFVCGGLAACMATL 134

RESULT 5

BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9305685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT "Analysis of maize brittle-1 alleles and a defective
RT suppressor-mutator-induced mutable allele";
RL Plant Cell 3:1337-1348(1991).
CC -1- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -1- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M79333; AAA33438.1; -.
CC PIR; JQ1459; JQ1459.
CC MaizeDB; 47578; -.
CC InterPro; IPR002067; Mit_carrier.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
FT CHAIN 76 436 BRITTLE-1 PROTEIN.
FT TRANSMEM 229 247 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
SQ SEQUENCE 436 AA; 46627 MW; 9600C05F603E9DAE CRC64;

Query Match 13.0%; Score 134.5; DB 1; Length 436;
Best Local Similarity 34.1%; Pred. No. 0.00058;
Matches 42; Conservative 16; Mismatches 52; Indels 13; Gaps 4;
QY 67 PAGATVTHMVAVAGVILEHCVMYIPDCVKTMRQSLQDPAPAAARNVLEALWRIIRTEG 126
Db 224 PPKPIPTPLVAGLAGFASTLCTYPMLIKTRV-TIEKDV---YDNVAFVKILRDEG 279
QY 127 LWRPMGLNVTATGAGPAHALYFACYEKLKTLSDV--HPGNSHTANGAACVATLLH 184
Db 280 PSELYRGLTSLIGVVPYAAACNFYAYETLKRRLYRRATGRRPGAD-----VGPVATLLI 332

QY 185 DAA 187
Db 333 GSA 335

RESULT 6

ODC_HUMAN STANDARD; PRT; 299 AA.
AC Q9BQ8;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 2-oxodicarboxylate carrier.
GN SLC25A21 OR ODC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RT TISSUE=Liver;
RX MEDLINE=21269385; PubMed=11083877;
RA Fiermonte G., Dolce V., Palmieri L., Ventura M., Runswick M.J.,
RA Palmieri F., Walker J.E.;
RT "Identification of the human mitochondrial oxodicarboxylate carrier:
RT bacterial expression, reconstitution, functional characterization,
RT tissue distribution and chromosomal location";
RL J. Biol. Chem. 276:8225-8230(2001).
CC -1- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC oxoglutarate, adipate, glutarate, and to a lesser extent,
CC pimelate, 2-oxopimelate, 2-aminoadipate, oxaloacetate, and
CC citrate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ278148; CAC27562.1; -.
CC InterPro; IPR001993; Mitoch_carrier.
CC InterPro; IPR002067; Mit_carrier.
CC Pfam; PF00153; mito_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 62 82 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 211 231 POTENTIAL.
SQ SEQUENCE 299 AA; 33303 MW; 69A259400328AE19 CRC64;

Query Match 12.9%; Score 134; DB 1; Length 299;
Best Local Similarity 31.7%; Pred. No. 0.00043;
Matches 38; Conservative 26; Mismatches 48; Indels 8; Gaps 4;
QY 76 MVAGVAGVILEHCVMYIPDCVKTMR--MQSLQDPAPAAARNVLEALWRIIRTEGLWRPMRG 133
Db 17 IVAGSGNLVEICMLHPDLVDVVKTRFQIQRCATDPNS-YKSLVDVSRFMFQMEGLFGYKG 75
QY 134 LNVATGAGPAHALYFACYEKLKTLSDV--HPGNSHTANGAACVATLLHDAAMPAPAE 192
Db 76 ILPPILAETPKRAVKFTTFQYKLLGVLSLSPALTPAIAGLSGLTEAIV---VNPFE 131

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF317711; AAG60687.1; -
 DR EMBL; AF151827; AAD34064.1; -
 DR EMBL; AL133584; CAB63728.1; -
 DR EMBL; AK026060; BAB15341.1; -
 DR EMBL; BC001398; AAH01398.1; -
 DR EMBL; BC009330; AAH09330.1; -
 DR EMBL; AF119864; AAF69618.1; -
 DR InterPro; IPR001993; Mitoch_carrier.
 DR InterPro; IPR002067; Mit_carrier.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Alternative splicing; Polymorphism.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 64 72 POTENTIAL.
 FT VARSPLIC 247 247 L -> F.
 FT VARIANT 247 247 /FTid=VAR_012756.
 FT CONFLICT 266 266 G -> R (IN REF. 6).
 FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

Query Match 12.3%; Score 127; DB 1; Length 359;
 Best Local Similarity 23.8%; Pred. No. 0.0019;
 Matches 45; Conservative 19; Mismatches 51; Indels 74; Gaps 5;

QY 67 PAGAVTTHWVAGAVAGILEHCVMPIDCVKTRMOSLQDPAA----- 109
 DB 6 PAGISPLQOMVASGTGAVVTVSLFMTPLDVVKVRLQSRPSMASELMPSSRLWSLSTKLP 65
 QY 110 -----RYRNVLEALWRIIRTEG---LWR 129
 DB 66 SLSQSTGKLLCYNGVLEPLILCPNGARCATWFDQPTRTGTMDAFVKIVRHEGFTLW- 124
 QY 130 PMRGLNVATGAGPAHALYFACYEKLKTL-----SDVIHPGGNSHIANGAGCVATLL 183
 DB 125 --SGLPATLVMTVPATATVFTAYDQKAFKLCGRALTSOLYAP-----WVAGALARLG 174
 QY 184 HDAAMNPAE 192
 DB 175 TVTVISPLE 183

RESULT 8
 TXTP_BOVIN
 ID TXTP_BOVIN STANDARD; PRT; 311 AA.
 AC P79110;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tricarboxylate transport protein, mitochondrial precursor (Citrate
 transport protein) (CTP) (Tricarboxylate carrier protein).
 GN SLC20A3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.

RESULT 7
 CG69_HUMAN
 ID CG69_HUMAN STANDARD; PRT; 359 AA.
 AC Q9BEJ4; Q9UF66; Q9Y379; Q9P182;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial carrier protein CGI-69 (PRO2163).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
 RP TISSUE SPECIFICITY.
 RC TISSUE=Liver;
 RX PubMed=11139402;
 RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
 RA Pan G., Adams S.H.;
 RT "Overexpression of the human 2-oxoglutarate carrier lowers
 RT mitochondrial membrane potential in HEK-293 cells: contrast with the
 RT unique cold-induced mitochondrial carrier CGI-69.";
 RL Biochem. J. 353:369-375(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20272150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 RT Caenorhabditis elegans by comparative proteomics.";
 RL Genome Res. 10:703-713(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ankerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Meyes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Colon;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC 1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC 1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
 CC are produced by alternative splicing.
 CC 1- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
 CC and kidney.
 CC 1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----

RX MEDLINE-97019278; PubMed-8865808;
 RA Iacobazzi V., de Palma A., Palmieri F.;
 RT "Cloning and sequencing of the bovine cDNA encoding the mitochondrial
 RL tricarboxylate carrier protein.";
 RL Biochim. Biophys. Acta 1284:9-12(1996).
 CC -!- FUNCTION: INVOLVED IN CITRATE-H+/MALATE EXCHANGE. IMPORTANT FOR
 CC THE BIOENERGETICS OF HEPATIC CELLS AS IT PROVIDES A CARBON SOURCE
 CC FOR FATTY ACID AND STEROL BIOSYNTHESIS, AND NAD+ FOR THE
 CC GLYCOLYTIC PATHWAY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X7773; CAA66375.1; -;
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_car; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 FT TRANSIT 1 13 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 14 311 TRICARBOXYLATE TRANSPORT PROTEIN.
 FT TRANSMEM 29 46 1 (POTENTIAL).
 FT TRANSMEM 86 105 2 (POTENTIAL).
 FT TRANSMEM 129 143 3 (POTENTIAL).
 FT TRANSMEM 183 202 4 (POTENTIAL).
 FT TRANSMEM 224 241 5 (POTENTIAL).
 FT TRANSMEM 278 297 6 (POTENTIAL).
 FT REPEAT 14 114 1.
 FT REPEAT 115 209 2.
 FT REPEAT 210 311 3.
 FT SEQUENCE 311 AA; 33876 MW; 75E010BC872F59BE CRC64;
 Query Match 12.1%; Score 125; DB 1; Length 311;
 Best Local Similarity 30.6%; Pred. No. 0.0024;
 Matches 30; Conservative 22; Mismatches 42; Indels 4; Gaps 1;
 QY 68 AGATVTTTH----MVAGAVAGILEHCVMYPIDCVKTRMSLODPDPAARYNVLEALWRIIR 123
 DB 17 SGKAKLTHPGKAILAGLAGGIEICITFTPEYVKTQLDERSHPPRYRGIGDCVQTVR 76
 QY 124 TGLWRPRLNVTATGAGPAHALYFACYEKLKLTSD 161
 DB 77 SHGLLGLRGLSLLYGSIPKAAVRGTFEFLSNHMRD 114

RESULT 9

ID TXTP_HUMAN STANDARD; PRT; 311 AA.
 AC P53007; Q9BSK6;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Tricarboxylate transport protein, mitochondrial precursor (Citrate
 DE transport protein) (CTP) (Tricarboxylate carrier protein).
 GN SLC25A1 OR SLC20A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96115597; PubMed-8666394;
 RA Heisterkamp N., Mulder M.P., Langeveld A., ten Hoeve J., Wang Z.,

RA Roe B., Groffen J.;
 RT "Localization of the human mitochondrial citrate transporter protein
 RL gene to chromosome 22q11 in the DiGeorge syndrome critical region.";
 RL Genomics 29:451-456(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN CITRATE-H+/MALATE EXCHANGE. IMPORTANT FOR
 CC THE BIOENERGETICS OF HEPATIC CELLS AS IT PROVIDES A CARBON SOURCE
 CC FOR FATTY ACID AND STEROL BIOSYNTHESIS, AND NAD+ FOR THE
 CC GLYCOLYTIC PATHWAY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U25147; AAB08515.1; -;
 DR EMBL: BC004980; AAH04980.1; -;
 DR EMBL: BC008061; AAH08061.1; -;
 DR MIM: 190315; -;
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_car; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 FT TRANSIT 1 13 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 14 311 TRICARBOXYLATE TRANSPORT PROTEIN.
 FT TRANSMEM 29 46 1 (POTENTIAL).
 FT TRANSMEM 86 105 2 (POTENTIAL).
 FT TRANSMEM 129 143 3 (POTENTIAL).
 FT TRANSMEM 183 202 4 (POTENTIAL).
 FT TRANSMEM 224 241 5 (POTENTIAL).
 FT TRANSMEM 278 297 6 (POTENTIAL).
 FT REPEAT 14 114 1.
 FT REPEAT 115 209 2.
 FT REPEAT 210 311 3.
 FT CONFLICT 26 26 G->E (IN REF. 1).
 FT SEQUENCE 311 AA; 34012 MW; F1341629924953D6 CRC64;
 Query Match 12.0%; Score 124.5; DB 1; Length 311;
 Best Local Similarity 31.4%; Pred. No. 0.0026;
 Matches 32; Conservative 22; Mismatches 43; Indels 5; Gaps 2;
 QY 65 ALPA-GATVTTTH----MVAGAVAGILEHCVMYPIDCVKTRMSLODPDPAARYNVLEALW 119
 DB 13 AAPASGKAKLTHPGKAILAGLAGGIEICITFTPEYVKTQLDERSHPPRYRGIGDCVR 72
 QY 120 RIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKLTSD 161
 DB 73 QTVRSHGVGLYGLSLLYGSIPKAAVRGTFEFLSNHMRD 114
 RESULT 10
 ID ORT2_HUMAN STANDARD; PRT; 301 AA.
 AC Q9BX12;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial ornithine transporter 2.
 GN ORNT2.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RX MEDLINE=21154914; PubMed=11230163;
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,
 RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;
 RT "Comparative DNA sequence analysis of mouse and human protocadherin
 RT gene clusters.";
 RL Genome Res. 11:389-404(2001).
 CC -!- FUNCTION: Ornithine transport across inner mitochondrial membrane,
 CC from the cytoplasm to the matrix (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AF332005; AAK26320.1; -.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 DR Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 KW TRANSMEM 5 25 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 110 130 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 237 257 POTENTIAL.
 SQ SEQUENCE 301 AA; 32580 MW; F22760BADE31DE21 CRC64;
 Query Match 11.8%; Score 122; DB 1; Length 301;
 Best Local Similarity 33.3%; Pred. No. 0.0041;
 Matches 28; Conservative 16; Mismatches 38; Indels 2; Gaps 1;
 QY 76 MVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYNRVLEALWRIIRTEGLWRPMTGLN 135
 Db 213 MLSGVAGICLWLWVFPVDCIKSRQVL--SMYKQAGFGITLLSVNRNEGIVALSGLK 270
 QY 136 VTATGAGPAHALYFACYEKLKLT 159
 Db 271 ATMIRAIPANGALFVAVEYSRKMM 294
 RESULT 11
 ODC1_YEAST STANDARD; PRT; 310 AA.
 AC OQ3028;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial 2-oxodicharboxylate carrier 1.
 GN ODC1 OR YPL134C OR LPI1LC.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
 RA Araujo R., Aparicio A., Barrell B., Badcock K., Benes V., Botstein D.,
 RA Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E.,
 RA Churcher C., Coster F., Davis K., Davis R.W., Dietrich F.S.,

RA Delius H., DiPaolo T., Dubois E., Dusterhoft A., Duncan M., Floeth M.,
 RA Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U.,
 RA Heumann K., Hilbert H., Hillier L., Hunnicke-Smith S., Hymen R., D.,
 RA Johnston M., Kaiman S., Kleine K., Komp C., Kurdi O., Lashkari F.,
 RA Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F.,
 RA Mewes H.W., Mirtipati S., Moestl D., Muller-Auer S., Namath A.,
 RA Purnelle U., Oefner P., Pearson D., Petel F.X., Pohl T.M.,
 RA Punnelle D., Schafer M., Scharfe M., Scherens B., Schramm S.,
 RA Schroeder M., Sdicu A.M., Tettelin H., Urrestarazu L.A., Ushinsky S.,
 RA Vierendeels F., Viissers S., Voss H., Walsh S.V., Wambutt R., Wang Y.,
 RA Wedler E., Wedler H., Winnett E., Zhong W.W., Zollner A., Vo D.H.,
 RA Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX PubMed=11013234;
 RA Palmieri L., Agrimi G., Runswick M.J., Fearnley I.M., Palmieri F.,
 RA Walker J.E.;
 RT "Identification in Saccharomyces cerevisiae of two isoforms of a novel
 RT mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate.";
 RL J. Biol. Chem. 276:1916-1922(2001).
 CC -!- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
 CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
 CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
 CC citrate and malate. The main physiological role is probably to
 CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
 CC matrix to the cytosol where they are used in the biosynthesis of
 CC lysine and glutamate, respectively, and in lysine catabolism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: U43703; AAB68225.1; -.
 DR SGD; S0006055; ODC1.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSMEM 9 29 POTENTIAL.
 FT TRANSMEM 126 146 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 310 AA; 34206 MW; 8EB7D631BC2BCEE CRC64;
 Query Match 11.7%; Score 121.5; DB 1; Length 310;
 Best Local Similarity 30.0%; Pred. No. 0.0046;
 Matches 36; Conservative 19; Mismatches 46; Indels 19; Gaps 4;
 QY 78 AGAVAGILEHCVMYPIDCVKTRMQ---SLQPDPA---RYNRVLEALWRIIRTEG 126
 Db 17 AGATAGVELLYMYPIDVVKTRMQLVTKGHPAAVAAKAADVHTGVMDCLTKVKG 76
 QY 127 LWRPMRGLNVTATGAGPAHALYFACYEKLKLTSLDVTHPGNSHI-----ANGACGV 179
 Db 77 FSHLYKGITSPILMEAPKRAIKFSGNDTF-QTYKKFFPTPNPGENGTOKIATYSGASAGV 135
 RESULT 12
 SFCL_YEAST STANDARD; PRT; 322 AA.
 ID SFCL_YEAST
 AC P33303;


```

RESULT 14
MFT_HUMAN
ID MFT_HUMAN STANDARD; PRT; 315 AA.
AC Q9H2D1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial folate transporter/carrier.
GN MFT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538421; PubMed=10979331;
RA Titus S.A.; Moran R.G.;
RT "Retrovirally mediated complementation of the glyB phenotype. Cloning
RT of a human gene encoding the carrier for entry of folates into
RT mitochondria."
RL J. Biol. Chem. 275:36811-36817(2000).
CC -!- FUNCTION: Transport folate across the inner membranes of
CC mitochondria.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF283645; AAC37834.1;
CC InterPro: IPR001993; Mitoch_carrier.
CC Pfam: PF00153; mito_carr; 3.
CC PRINTS: PR00926; MITOCARRIER.
CC PROSITE: PS00215; MITOCH_CARRIER; 1.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 89 106 POTENTIAL.
FT TRANSMEM 227 243 POTENTIAL.
FT TRANSMEM 281 300 POTENTIAL.
SQ SEQUENCE 315 AA; 35388 MW; 60EBC0D61951EB6E CRC64;

Query Match 11.5%; Score 119.5; DB 1; Length 315;
Best Local Similarity 27.5%; Pred. No. 0.0068;
Matches 41; Conservative 23; Mismatches 54; Indels 31; Gaps 7;

QY 68 AGATVTH-----WVAGVAGILEHCVMYPIDCVKTRMQ-----SIOPDPAARYNRVLEA 117
DB 13 AWSTVFRHVRVYENLGIAGVSGVLSNLALHPLDLVLRFAVSDGLELRP-----KYNGLHC 68

QY 118 LWRIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLK--KTLSDVIHPGNGNSHIANGA 175
DB 69 LTTINKDLGLRGVLPNGVINGAGLSWGLYFFFNNAIKSYKTEGRAHLEATEYLVSA 128

QY 176 -AG-----CVAT-----LLHDAAMN 189
DB 129 EAGATMLCITNPLWVTKTRMLQYDAVYN 157

RESULT 15
ADT_ANOGA
ID ADT_ANOGA STANDARD; PRT; 301 AA.
AC Q27238;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ADP/ATP carrier protein (ADP/ATP translocase) (Adenine nucleotide
DE translocator) (ANT).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G3;
RX MEDLINE=94348635; PubMed=8069414;
RA Beard C.B.; Crews-Oyen A.E.; Kumar V.K.; Collins F.H.;
RT "A cDNA encoding an ADP/ATP carrier from the mosquito Anopheles
RT gambiae."
RL Insect Mol. Biol. 3:35-40(1994).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
CC MITOCHONDRIAL INNER MEMBRANE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L11618; AAB04104.1;
CC EMBL: L11617; AAB04105.1;
CC InterPro: IPR002067; Mit_carr.
CC InterPro: IPR001993; Mitoch_carrier.
CC Pfam: PF00153; mito_carr; 3.
CC PRINTS: PR00926; MITOCARRIER.
CC PROSITE: PS00215; MITOCH_CARRIER; 3.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 14 31 1 (POTENTIAL).
FT TRANSMEM 75 93 2 (POTENTIAL).
FT TRANSMEM 119 136 3 (POTENTIAL).
FT TRANSMEM 178 197 4 (POTENTIAL).
FT TRANSMEM 216 233 5 (POTENTIAL).
FT TRANSMEM 275 293 6 (POTENTIAL).
SQ SEQUENCE 301 AA; 32863 MW; 4CC9E17C9F8DA08B CRC64;

Query Match 11.4%; Score 118.5; DB 1; Length 301;
Best Local Similarity 29.2%; Pred. No. 0.0078;
Matches 31; Conservative 19; Mismatches 51; Indels 5; Gaps 3;

QY 78 AGAVAGILEHCVMYPIDCVKTRM-QSLOPDPAAR-YRNVLEALWRIIRTEGLWRPMRGLN 135
DB 121 SGAGAGATSLCFVYPIDLFARTRLGADVGPAGERFNGLLDCLKTKVSDGIIGLYRGN 180

QY 136 VTATGAGPAHALYFACYEKLKKTLSVDVIHPGNGNSHIANGAAGCVAT 181
DB 181 VSVQGIIVYRAAYFCGDFATKGLPD---PKNTSIFVSWAIQVVT 223

Search completed: August 27, 2002, 04:16:46
Job time: 1238 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:20 ; Search time 130.92 Seconds
(without alignments)
255.026 Million cell updates/sec

Title: US-09-870-113-4

Perfect score: 1036

Sequence: 1 MELEGRGAGGVAGGPAAGPG.....GAAGCVATLLHDAAMNPAEG 193

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1030	99.4	364	4 Q96A46	Q96A46 homo sapien
2	516.5	49.9	338	11 Q920G8	Q920G8 mus musculus
3	515.5	49.8	347	4 Q9NYZ2	Q9NYZ2 homo sapien
4	444.5	42.9	155	4 Q969S1	Q969S1 homo sapien
5	444.5	42.9	182	11 Q91ZY0	Q91ZY0 mus musculus
6	338.5	32.7	312	5 Q231Z5	Q231Z5 caenorhabdi
7	295	28.5	379	5 Q9VAY3	Q9VAY3 drosophila
8	291	28.1	380	5 Q9NHV6	Q9NHV6 drosophila
9	267.5	25.8	303	5 Q94638	Q94638 onchocerca
10	265.5	25.6	301	5 Q94634	Q94634 onchocerca
11	252	24.3	303	3 Q14281	Q14281 schizosacch
12	224.5	21.7	781	10 Q9LWJ6	Q9LWJ6 arabidopsis
13	223	21.5	331	10 Q64731	Q64731 arabidopsis
14	220	21.4	289	10 Q82049	Q82049 ribes nigru
15	158.5	15.3	147	11 Q9D547	Q9D547 mus musculus
16	155.5	15.0	322	10 Q9AX03	Q9AX03 oryza sativ

17	150	14.5	310	5 Q18844	Q18844 caenorhabdi
18	144.5	13.9	320	11 Q921P8	Q921P8 mus musculus
19	144.5	13.9	399	5 Q9VKZ5	Q9VKZ5 drosophila
20	143.5	13.9	449	5 Q9VI05	Q9VI05 drosophila
21	143	13.8	358	10 Q9M333	Q9M333 arabidopsis
22	142	13.7	348	10 Q9SH98	Q9SH98 arabidopsis
23	142	13.7	413	5 Q9VWF9	Q9VWF9 drosophila
24	139	13.4	298	11 Q99JD3	Q99JD3 rattus norv
25	137.5	13.3	321	4 Q9BSK2	Q9BSK2 homo sapien
26	136.5	13.2	304	4 Q9NWX2	Q9NWX2 homo sapien
27	136.5	13.2	304	4 Q96DW6	Q96DW6 homo sapien
28	136.5	13.2	334	11 Q9CW38	Q9CW38 mus musculus
29	136	13.1	319	10 Q9MA27	Q9MA27 arabidopsis
30	135.5	13.1	320	4 Q9HC21	Q9HC21 homo sapien
31	134.5	13.0	345	11 Q9CYJ1	Q9CYJ1 mus musculus
32	134	12.9	299	4 Q9BQT8	Q9BQT8 homo sapien
33	133.5	12.9	322	10 Q9ARL9	Q9ARL9 hordeum vul
34	133	12.8	361	3 Q9P6L7	Q9P6L7 schizosacch
35	132.5	12.8	291	3 Q13844	Q13844 schizosacch
36	131.5	12.7	297	5 Q9VBN7	Q9VBN7 drosophila
37	131.5	12.7	317	5 Q9VGM3	Q9VGM3 drosophila
38	131	12.6	310	4 Q9NVN5	Q9NVN5 homo sapien
39	131	12.6	311	4 Q96CQ1	Q96CQ1 homo sapien
40	131	12.6	311	11 Q922G0	Q922G0 mus musculus
41	131	12.6	351	4 Q9Y379	Q9Y379 homo sapien
42	131	12.6	351	4 Q9UF66	Q9UF66 homo sapien
43	130	12.5	305	10 Q9ZWG1	Q9ZWG1 arabidopsis
44	129	12.5	313	10 Q9SUY5	Q9SUY5 arabidopsis
45	128.5	12.4	306	5 Q9V3T2	Q9V3T2 drosophila

ALIGNMENTS

RESULT 1

Q96A46	PRELIMINARY;	PRT;	364 AA.
ID Q96A46			
AC Q96A46;			
DT 01-DEC-2001 (TREMBLrel. 19, Created)			
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DE MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE			
DE CARRIER SPLICE VARIANT).			
GN HMRS3/4			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-21195335; PubMed-11297739;			
RA Li F., Nikali K., Grogan J., Leibiger I., Leibiger B., Schweyen R.,			
RA Larsson C., Suomalainen A.;			
RT "Characterization of a novel human putative mitochondrial transporter			
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";			
RL FEBS Lett. 494:79-84(2001).			
DR EMBL; AJ303077; CAC27996.1; -			
DR EMBL; AF327402; AAK49519.1; -			
SQ SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;			

Query Match 99.4%; Score 1030; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.3e-82;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELEGRGAGGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACACRPVRODPDSG 60

Db 1 MELEGRGAGGVAGGPAAGPGRSPGESALLDGLQVGRGAGGAGACACRPVRODPDSG 60

QY 61 PDVEALPAGATVTTTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVEALWR 120

Db 61 PDVEALPAGATVTTTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPAARYRNVEALWR 120

QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNSHIANGAAGCVA 180
DB 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNSHIANGAAGCVA 180
QY 181 TLLHDAAMNPAE 192
DB 181 TLLHDAAMNPAE 192
RESULT 2
Q920G8 PRELIMINARY; PRT; 338 AA.
AC Q920G8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
RA She J.-X.;
RT "A new gene which is highly expressed in NOD mice spleen."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF288621; AAL23859.1; -
SQ SEQUENCE 338 AA; 37310 MW; BB35B1F70C56A3FE CRC64;

Query Match 49.9%; Score 516.5; DB 11; Length 338;
Best Local Similarity 54.7%; Pred. No. 2e-37;
Matches 105; Conservative 19; Mismatches 41; Indels 27; Gaps 3;
QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
DB 1 MELRRGVGVGNOAAG-----RRMDGDCRDG--GCGSKDAGS----- 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
DB 34 EDYENLPTSASVSTHTAGAMAGILEHSVMYPVDSVKTRMQSLSDPKRAQYTSIYGALKR 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNSHIANGAAGCVA 180
DB 94 IMHTEGFWRPLRGVNMVMGAGPAHAMVFACYENKMKRTLNDVFSHQGNSHLANGVAGSVA 153
QY 181 TLLHDAAMNPAE 192
DB 154 TLLHDAAMNPAE 165

RESULT 3
Q9NYZ2 PRELIMINARY; PRT; 347 AA.
AC Q9NYZ2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-HYPOTHALAMUS;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF223466; AAF64141.1; -
DR InterPro: IPR001993; Mitoch_carrier.

DR Pfam: PF00153; mito_carri; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWNL.2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 49.8%; Score 515.5; DB 4; Length 347;
Best Local Similarity 55.2%; Pred. No. 2.5e-37;
Matches 106; Conservative 20; Mismatches 39; Indels 27; Gaps 4;
QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
DB 1 MELR---SGSV--GSQAVARRMDGDS-----RDGGG-----KDATGS 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
DB 34 EDYENLPTSASVSTHTAGAMAGILEHSVMYPVDSVKTRMQSLSDPKRAQYTSIYGALKK 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNSHIANGAAGCVA 180
DB 94 IMRTGEGFWRPLRGVNMVMGAGPAHAMVFACYENKMKRTLNDVFFHQGNSHLANGIAGSNA 153
QY 181 TLLHDAAMNPAE 192
DB 154 TLLHDAAMNPAE 165

RESULT 4
Q969S1 PRELIMINARY; PRT; 155 AA.
AC Q969S1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITOCHONDRIA SOLUTE CARRIER PROTEIN (HYPOTHETICAL 16.8 KDA
DE PROTEIN).
GN MSCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Q., Eckenrode S., Ruan Q., Wang C., Shi J., McIndoe R.A., She J.;
RT "Molecular cloning of a novel mitochondria solute carrier protein
RT (MSCP) gene from mouse and human and its down-regulation in mouse
RT spleen during the maturation of the immune system."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-UTERUS, AND LEIOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY032628; AAK38154.1; -
DR EMBL: BC015013; AAH15013.1; -
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16832 MW; 8ACB98A483C8E6EF CRC64;

Query Match 42.9%; Score 444.5; DB 4; Length 155;
Best Local Similarity 52.9%; Pred. No. 1.6e-31;
Matches 92; Conservative 19; Mismatches 36; Indels 27; Gaps 4;
QY 1 MELEGRGAGGAGGPAAGPGRSPGESALLDGLQRGVGRGAGGAGACRPPVRQDPDSG 60
DB 1 MELR---SGSV--GSQAVARRMDGDS-----RDGGG-----KDATGS 33
QY 61 PDYEALPAGATVTHWAGAVAGILEHCVMYPIDCVKTRMQSLQDPDPAARYNRVLEALWR 120
DB 34 EDYENLPTSASVSTHTAGAMAGILEHSVMYPVDSVKTRMQSLSDPKRAQYTSIYGALKK 93
QY 121 IIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNSHIANG 174
DB 94 IMRTGEGFWRPLRGVNMVMGAGPAHAMVFACYENKMKRTLNDVFFHQGNSHLANG 147

Query Match	32.7%;	Score 338.5;	DB 5;	Length 312;	
Best Local Similarity	48.0%;	Pred. No. 7e-22;			
Matches	73;	Conservative 23;	Mismatches 37;	Indels 19;	Gaps 4;
QY	42	GGGEAGACRPVVRQDPDSDGYEALPAGATVTHWVAGAVAGIIEHCVMPIDCVKTRMQ	101		
Db	3	GGGE-----DEYESLPT-HSVPVHLTAGALAGAVEHCVMPFDSVTRMQ	46		
QY	102	SLQPDPAARYRNVLALWRIIRTEGLWRPRLGLNVTATGAGPAHALYFACYEKLKTKLSD	161		
Db	47	SLQCPDPTKCPPTPVHSLMSIVKREGWLRPLRGVNAAGSMPAHALYFTVYEKKMGYLTG	106		
QY	162	VTHPGGNSH-IANGAGACVATLLHDAAMPAE	192		
Db	107	--NSAGHSNTLAYGASGVVATLIHDAIMPAE	136		
RESULT	7				
Q9VAY3		PRELIMINARY;	PRT;	379 AA.	
ID	Q9VAY3;				
AC	Q9VAY3;				
DT	01-MAY-2000 (TREMBlrel. 13, Created)				
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)				
DE	CG4963 PROTEIN (GH09840P).				
GN	CG4963.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BERKELEY;				
RC	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,				
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,				
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,				
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,				
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,				
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,				
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,				
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,				
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,				
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;				
RT	"The genome sequence of Drosophila melanogaster.";				
RL	Science 287:2185-2195(2000).				

Query Match	42.9%;	Score 444.5;	DB 11;	Length 182;	
Best Local Similarity	52.3%;	Pred. No. 2e-31;			
Matches	91;	Conservative 18;	Mismatches 38;	Indels 27;	Gaps 3;
QY	1	MELEGRGAGVAGGPAAGPGRSPGESALLDGLQGVGRGAGGGAGACGPPVRQDPDSG	60		
Db	1	MELRRGGVGNQAAG-----RRMDGDCRDG---CGGSKDAGS-----	33		
QY	61	PDYEALPAGATVTHWVAGAVAGIIEHCVMPIDCVKTRMQSLQDPDPAARYRNVLALWR	120		
Db	34	EDYENLPTASVSTHMTAGAGIIEHIMVVDVSKTRMQSLNPDPKARYTSIYGALKR	93		
QY	121	IIRTEGLWRPRLGLNVTATGAGPAHALYFACYEKLKTKLSDVIHPGGNSHIANG	174		
Db	94	IMHTGCFWRPLRLGLNMMGAGPAHAFYACVYENKMKRLINDVFSHQNSHLANG	147		
RESULT	6				
Q23125		PRELIMINARY;	PRT;	312 AA.	
ID	Q23125;				
AC	Q23125;				
DT	01-NOV-1996 (TREMBlrel. 01, Created)				
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)				
DE	01-DEC-2001 (TREMBlrel. 19, Last annotation update)				
DE	W02B12.9 PROTEIN.				
GN	W02B12.9.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;				
OC	Rhabditidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	SwInburne J., Ainscough R.;				
RL	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	MEDLINE=99069613; PubMed=9851916;				
RA	none;				
RT	"Genome sequence of the nematode C.elegans: A platform for				
RT	investigating biology.";				
RL	Science 282:2012-2018(1998).				
DR	EMBL; Z66521; CAA91399.1; --				
DR	InterPro; IPR001993; Mitoch_carrier.				
DR	Pfam; PF00153; mito_carr; 3.				
DR	PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.				
SQ	SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;				

```

[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Erise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AE003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carri; 3.
DR PRINTS; PS00926; MITOCHARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 28.5%; Score 295; DB 5; Length 379;
Best Local Similarity 47.3%; Pred. No. 5.6e-18;
Matches 62; Conservative 22; Mismatches 41; Indels 6; Gaps 3;

QY 62 DYELPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPAPAAARNVLEALWRI 121
DB 5 DYESLPT-TSVGVMNTAGAGVLEHVMPPLDSVTRMQSL--SPPTKMNIVSTLRTM 61
QY 122 IRTEGLRPMRGUNLVNTATGAGPAHALYFCYCEKTKTSLSDVIHPGNSHIANGAAGCVAT 181
DB 62 ITREGLLRPIRGASAVVLGAGPASHLYFRAYEMTKELTAKFTSVRNLYVISGA---VAT 118

QY 182 LLHDAAMNPAE 192
DB 119 LIHDAISSPTD 129

RESULT 8
Q9NH96 ID Q9NH96 PRELIMINARY; PRT; 380 AA.
AC Q9NH96;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carri; 3.
DR PRINTS; PS00926; MITOCHARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 28.1%; Score 291; DB 5; Length 380;

```

```

Best Local Similarity 46.6%; Pred. No. 1.3e-17;
Matches 61; Conservative 22; Mismatches 42; Indels 6; Gaps 3;

QY 62 DYELPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPAPAAARNVLEALWRI 121
DB 5 DYESLPT-TSVGVMNTAGAGVLEHVMPPLDSVTRMQSL--SPPTKMNIVSTLRTM 61
QY 122 IRTEGLRPMRGUNLVNTATGAGPAHALYFCYCEKTKTSLSDVIHPGNSHIANGAAGCVAT 181
DB 62 ITREGLLRPIRGASAVVLGAGPASHLYFRAYEMTKELTAKFTSVRNLYVISGA---VAT 118

QY 182 LLHDAAMNPAE 192
DB 119 LIHDAISSPTD 129

RESULT 9
Q94638 ID Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carri; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 25.8%; Score 267.5; DB 5; Length 303;
Best Local Similarity 37.5%; Pred. No. 1.1e-15;
Matches 54; Conservative 26; Mismatches 41; Indels 23; Gaps 2;

QY 49 CRPPVRQDPDSDYEALPAGATVTTTHMVAGAVAGILEHCVMPIDCVKTRMQSLQDPDA 108
DB 12 CRWPV-----HLLAGSVAGLAELHCLMFPDPSVKTRLOSCLPCPE 50
QY 109 ARYRNVLALWRIIRTEGLRPMRGUNLVNTATGAGPAHALYFCYCEKTKTSLSDVIHPGNN 168
DB 51 TSCPTAMHSLMSVKKREGLLSLKGVNAVVLGTPAHAFYTYVYENSKAYLLN--NPRVS 108
QY 169 SHIANGAAGCVATLLHDAAMNPAE 192
DB 109 NSVSYAISGALATVIHDAVMNPAE 132

RESULT 10
Q94634 ID Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=96326580; PubMed=8703971;
RA CaMull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RT mitochondrial solute carriers.";
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45997; AAB19036.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; COBA8D819FB8EA79 CRC64;

Query Match 25.6%; Score 265.5; DB 5; Length 301;
Best Local Similarity 37.5%; Pred. No. 1.7e-15;
Matches 54; Conservative 26; Mismatches 41; Indels 23; Gaps 2;

Qy 49 CRPPVQPDSPDYALPAGATVTHMVAGAGILEHCWMPIDCVKTRMQSLQPDPA 108
Db 12 CRWPV-----HLLAGSVAGLAELHCLMFPDPSVKTRLSLCPCE 50

Qy 109 ARYRNVLALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKTTLSVDVTHPGGN 168
Db 51 TSCPAMHSLMSWKREGLLSKGVNAVLTGTPAHAFYTVYENSXAYLLN--NPRVS 108

Qy 169 SHIANGAAGCVATLLHDAAMNPAE 192
Db 109 NSMSYAISGALATVTHDAVMNPAE 132

RESULT 11
O14281 PRELIMINARY; PRT; 303 AA.
AC O14281;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER C8C9.12C.
GN SPAC8C9.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; 299168; CAB16300.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
SQ SEQUENCE 303 AA; 32652 MW; 65975CDE18107AB5 CRC64;

Query Match 24.3%; Score 252; DB 3; Length 303;
Best Local Similarity 40.3%; Pred. No. 2.5e-14;
Matches 52; Conservative 25; Mismatches 48; Indels 4; Gaps 1;

Qy 62 DYALPAGATVTHMVAGAGILEHCWMPIDCVKTRMQSLQPDPAARYRNVLALWRI 121
Db 9 DYEGLPIGPMYAHLLAGAFSGILEHSVMYPVDAIKTRMQLNGVSRVSGNIVNSVIKI 68
```

```

Qy 122 IRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKTTLSVDVTHPGNSHIANGAAGCVAT 181
Db 69 SSTEGLISWRGISSVIMGAGPSHAIYFSVLEFFSK-----INASDPRPLASALAGACAI 124

Qy 182 LLDHDAAMNP 190
Db 125 TISDAFMTP 133

RESULT 12
O9LMJ6 PRELIMINARY; PRT; 781 AA.
AC O9LMJ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC067971; AAF82217.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 4.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 781 AA; 87081 MW; 9FB579B9BD746D1E CRC64;

Query Match 21.7%; Score 224.5; DB 10; Length 781;
Best Local Similarity 38.0%; Pred. No. 1.9e-11;
Matches 57; Conservative 23; Mismatches 55; Indels 15; Gaps 5;

Qy 52 PVRQPD-----SGPDY--EALPAGATVTHMVAGAGILEHCWMPIDCVKTRMQSLQ 104
Db 9 PKFQEPDLRQVSQTPDFKPEIAHDLGLKFWQFIAGSVGVERHAMFPVDITKTHMQALR 68

Qy 105 PDPAARYRNVLALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKTTLSVDVTH 164
Db 69 PCP-LKPVGIREAFRSIIQKEGFSALYRGIWAMGLGAGPAHAHVYFSFVEVSKKYL---- 123

Qy 165 PGG--NSHIANGAAGCVATLLHDAAMNPAE 192
Db 124 -AGDQNSVAMSGVFAITSSDAVFTPM 152

RESULT 13
O64731 PRELIMINARY; PRT; 331 AA.
AC O64731;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER PROTEIN.
GN AT2G30160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
```


Db 73 EDYENLPTSASVSTHMTAGAMAGILEHSIMYPVDSVK 109
||| || | : | : ||| ||| : ||| : ||| : |||

Search completed: August 27, 2002, 04:19:22
Job time: 1289 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:58 ; Search time 138.54 Seconds
(without alignments)
212.463 Million cell updates/sec

Title: US-09-870-113-8

Perfect score: 1406

Sequence: 1 MQSLQDPAAARYNVLALW.....VYEFKYLITKQEWRAKG 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1406	100.0	289	21 AAB50383	Human uncoupling p
2	1406	100.0	289	22 AAB60113	Human transport pr
3	1406	100.0	318	22 AAM41505	Human polypeptide
4	1377	97.9	268	22 AAM9719	Human polypeptide
5	1013.5	72.1	272	21 AAB42980	Human ORFX ORF2744
6	999.5	71.1	366	22 AAM00938	Human bone marrow
7	916	65.1	174	22 AAU01989	Human secreted pro
8	910	64.7	176	21 AAB40634	Human ORFX ORF398
9	910	64.7	176	22 AAB98221	Human mitochondria
10	848.5	60.3	331	22 AAB60658	Human mitochondria
11	754.5	53.7	677	22 AAU29748	Novel human secret

12	739.5	52.6	199	22	AAM80023	Human protein SEQ
13	725.5	51.6	187	22	AAM93909	Human polypeptide,
14	725.5	51.6	187	22	AAB98222	Human mitochondria
15	725.5	51.6	187	22	AAB49665	Human mitochondria
16	717.5	51.0	187	22	AAM93934	Human polypeptide,
17	717.5	51.0	188	21	AAB32112	Human secreted pro
18	692	49.2	299	22	AAM79039	Human protein SEQ
19	657.5	46.8	379	22	ABB61130	Drosophila melanog
20	558	39.7	215	22	AAB49666	O. volvulus mitoch
21	541	38.5	194	21	AAB32111	Human secreted pro
22	515.5	36.7	181	22	AAU01988	Gene #25 human sec
23	429.5	30.5	289	18	AAW17054	Blackcurrant RIB7
24	416	29.6	278	21	AAG22079	Arabidopsis thalia
25	416	29.6	290	21	AAG22078	Arabidopsis thalia
26	416	29.6	331	21	AAG22077	Arabidopsis thalia
27	414	29.4	278	21	AAG43094	Arabidopsis thalia
28	414	29.4	290	21	AAG43093	Arabidopsis thalia
29	414	29.4	331	21	AAG43092	Arabidopsis thalia
30	412	29.3	111	21	AAB42968	Human ORFX ORF2732
31	343.5	24.4	149	21	AAB42964	Human ORFX ORF2728
32	309	22.0	74	22	ABG01167	Novel human diagno
33	276	19.6	155	21	AAB42966	Human ORFX ORF2730
34	266	18.9	413	22	ABB69108	Drosophila melanog
35	261.5	18.6	358	22	ABB71142	Drosophila melanog
36	251.5	17.9	351	21	AAB50378	Human uncoupling p
37	243.5	17.3	365	21	AAB42783	Human ORFX ORF2547
38	242.5	17.2	301	21	AAG52781	Arabidopsis thalia
39	242.5	17.2	301	21	AAG52783	Arabidopsis thalia
40	242.5	17.2	325	21	AAG52780	Arabidopsis thalia
41	242.5	17.2	325	21	AAG52782	Arabidopsis thalia
42	242.5	17.2	361	21	AAG52779	Arabidopsis thalia
43	237	16.9	449	22	ABB59263	Drosophila melanog
44	233	16.6	296	21	AAB43372	Human ORFX ORF3136
45	232	16.5	374	22	AAM40959	Human polypeptide

ALIGNMENTS

RESULT 1

AAB50383
ID AAB50383 standard; Protein; 289 AA.

XX AC AAB50383;

XX DT 12-MAR-2001 (first entry)

XX DE Human uncoupling protein #6.

XX KW Human; uncoupling protein; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cardiant; vasotropic;

KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;

KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

XX infertility.

XX KW Homo sapiens.

XX OS WO2000061614-A2.

XX PN 19-OCT-2000.

XX PD 06-APR-2000; 2000WO-US09534.

XX PF 09-APR-1999; 99US-0128701.

XX PR 08-JUL-1999; 99US-0142821.

XX PR 18-AUG-1999; 99US-0149448.

XX PR 12-NOV-1999; 99US-0164751.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

XX

DR WPI; 2000-656322/63.
 DR N-PSDB; AAC90457.
 XX
 XX Uncoupling proteins and nucleic acid sequences encoding them, useful
 PT for detecting, preventing and treating proliferative, neurological,
 PT immune system, cardiovascular and gastrointestinal disorders -
 XX
 XX Claim 11; Page 323-324; 343pp; English.
 PS
 XX The present sequence is a human uncoupling protein. The nucleotide
 CC sequences encoding the uncoupling proteins may be used for
 CC the detection of various disorders such as cancer, for chromosome
 CC identification, as chromosome markers and for numerous other diagnostic
 CC or research purposes. The uncoupling protein encoded by the nucleotide
 CC sequences may be used to treat disorders such as neural, immune,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders, wounds, infectious diseases,
 CC thrombosis, arthritis, and infertility.
 XX
 XX Sequence 289 AA;
 SQ

Query Match 100.0%; Score 1406; DB 21; Length 289;
 Best Local Similarity 100.0%; Pred. No. 3.1e-153;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOSLQDPAAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLL 60
 DB 25 mqslqpdpaaarynvlealwriirteglwrpmlgnvtatgagpahalyfacyekllkktl 84
 QY 61 SDVIHPGNSHIANGAAGCVATLLHDAAMPAAEVVKQRMQMYNSPYHRVDCVRAVWQNE 120
 DB 85 sdvihpgnshiangagcvatllhdaampaevvkqrmqmyspyhrvdcvrvwqne 144
 QY 121 GAGAFYRSYTTQLTMNVPPFOAIHPMTVEFLQEHFNPORRNPSSHVLSGACAGAVAAAT 180
 DB 145 gagafyrsyttqltmnvppfoaihpmfveflqehfnpryppsshvlsagacagavaaaat 204
 QY 181 TPLDVCVKTLLNTQESLALNSHITGHITGMASAFRTVYQVGVTAAYFRGVQARVIYQIPST 240
 DB 205 tpldvcvktllntqeslalnshitghitgmasafrtvyqvgvtaayfrgvqarvlyqipst 264
 QY 241 AIAMSVYEFFKYLITKRQEWRAK 265
 DB 265 aiawsvyeffkylitkrqeewrak 289

RESULT 2
 AAB60113
 ID AAB60113 standard; Protein; 289 AA.
 XX
 AC AAB60113;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human transport protein TPPT-33.
 XX
 KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200078953-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16668.
 XX
 PR 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.
 PR 28-OCT-1999; 99US-0162287.

XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
 XX WPI; 2001-041124/05.
 DR N-PSDB; AAF27733.
 XX
 XX Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX
 XX Claim 2; Page 133-134; 165pp; English.
 XX
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 XX
 XX Sequence 289 AA;
 SQ

Query Match 100.0%; Score 1406; DB 22; Length 289;
 Best Local Similarity 100.0%; Pred. No. 3.1e-153;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOSLQDPAAARYNVLEALWRIIRTEGLWRPMLGNVTATGAGPAHALYFACYEKLKKTLL 60
 DB 25 mqslqpdpaaarynvlealwriirteglwrpmlgnvtatgagpahalyfacyekllkktl 84
 QY 61 SDVIHPGNSHIANGAAGCVATLLHDAAMPAAEVVKQRMQMYNSPYHRVDCVRAVWQNE 120
 DB 85 sdvihpgnshiangagcvatllhdaampaevvkqrmqmyspyhrvdcvrvwqne 144
 QY 121 GAGAFYRSYTTQLTMNVPPFOAIHPMTVEFLQEHFNPORRNPSSHVLSGACAGAVAAAT 180
 DB 145 gagafyrsyttqltmnvppfoaihpmfveflqehfnpryppsshvlsagacagavaaaat 204
 QY 181 TPLDVCVKTLLNTQESLALNSHITGHITGMASAFRTVYQVGVTAAYFRGVQARVIYQIPST 240
 DB 205 tpldvcvktllntqeslalnshitghitgmasafrtvyqvgvtaayfrgvqarvlyqipst 264
 QY 241 AIAMSVYEFFKYLITKRQEWRAK 265
 DB 265 aiawsvyeffkylitkrqeewrak 289

RESULT 3
 AAM41505
 ID AAM41505 standard; Protein; 318 AA.
 XX
 AC AAM41505;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6436.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.

```
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 03-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI60661.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6436; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 318 AA;
XX
XX Query Match 100.0%; Score 1406; DB 22; Length 318;
XX Best Local Similarity 100.0%; Pred. No. 3 6e-153;
XX Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQSLQDPDPAARYRNVLBALWRIIRTEGLWRPMLGNTATGAGPAHALYFACYEKLKTL 60
Db |||||
54 mqslqdpdpaaryrnvlbalwriirteglwrpmlgntatgagpahalyfacyekllkttl 113
QY 61 SDVIHPGCGNSHIANGAGCVATLLHDAAMPNPAEVVKORMQYNPSYHRVTDCAVAVQNE 120
Db |||||
114 sdvihpgcgnshiangaagcvatllhdaampnaevvkrmqmynpsyhrvtcdvavvqne 173
QY 121 GAGAFYSYTTQTMNVFQAIHMTVEFLQEHFNPPORRYNPSHVLGSGACAGAVAAAT 180
Db |||||
174 gagafysyttqtmnvfqaiahmtveflqehfnpporrynpsshvlgsgacagavaaat 233
QY 181 TPLDVCKTLNTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVQARVIYQIPST 240
Db |||||
234 tpldvcktlntqeslalnshigtgmasaftrtyvqgvgtayfrgvqarviyqipst 293
QY 241 AIAWSYVEFFKYLITRKQEWRAK 265
Db |||||
294 aiawsyveffkylitkrqewragk 318
XX
RESULT 4
AAM39719
ID AAM39719 standard; Protein: 268 AA.
XX
```

```
AC AAM39719;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 2864.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX N-PSDB; AAI58875.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 4; SEQ ID NO 2864; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 268 AA;
XX
XX Query Match 97.9%; Score 1377; DB 22; Length 268;
XX Best Local Similarity 98.1%; Pred. No. 6.2e-150;
XX Matches 260; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 1 MQSLQDPDPAARYRNVLBALWRIIRTEGLWRPMLGNTATGAGPAHALYFACYEKLKTL 60
Db |||||
4 mslqdpdpaaryrnvlbalwriirteglwrpmlgntatgagpahalyfacyekllkttl 63
QY 61 SDVIHPGCGNSHIANGAGCVATLLHDAAMPNPAEVVKORMQYNPSYHRVTDCAVAVQNE 120
```

Db 64 sdvllhpggnshlangagcvatlhlhaanpaevvkqrmqmynsphrvtdcvravwqne 123
QY 121 GAGAFYRSYTTQLTMNVPQALHFMTYEFQLQEHFNPPRRYNPSSHVLGACAGAVAAAAT 180
Db 124 gagafysyrtqltmvnpqalhfntefyqlqehfnpprrynpsshvlsgacagavaaaat 183
QY 181 TPLDVCKTLNLTQESLALNSHTGHTGMSAFRTVYQGVGTAYPRGQVQARVIYQIPST 240
Db 184 tpldvcktlntqeslalsnshthgthgmsaftrvyqgvgtayfgrvgarviyqipst 243
QY 241 ATAWSVEFFKYLITKROEWRAK 265
Db 244 atawsveffkylitkrgewragk 268

RESULT 5
AAB42980
ID AAB42980 standard; Protein; 272 AA.
XX
AC AAB42980;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORP2744 polypeptide sequence SEQ ID NO:5488.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
PI WPI; 2000-602362/57.
XX N-PSDB; AAC77189.
DR
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 4662-4663; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 272 AA;
Query Match 72.1%; Score 1013.5; DB 21; Length 272;
Best Local Similarity 69.7%; Pred. No. 4.9e-108;
Matches 184; Conservative 39; Mismatches 40; Indels 1; Gaps 1;
QY 1 MOSLPDPAARYRNVLALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTYL 60
Db 7 mqslspdpkaqysiygalkimrtgfwrlrgvnmimgagpahamyfacyenmkrtl 66
QY 61 SDVlHPGGSHTANGAGCVATLLHDAAMNPAFVVKQRMQYNSPHRYTDCVRAVWQNE 120
Db 67 ndvlfhqqnshlangagsmatllhdavmnpaevvkqrlqmynsqhsaicsrtvwite 126
QY 121 GAGAFYRSYTTQLTMNVPQALHFMTYEFQLQEHFNPPRRYNPSSHVLGACAGAVAAAAT 180
Db 127 glgafysyrtqltmnlpqshlfityeflqeqvnphtynpqshtisgglagaaaaat 186
QY 181 TPLDVCKTLNLTQESLALN-SHTGHTGMSAFRTVYQGVGTAYFRGQVQARVIYQIPS 239
Db 187 tpldvcktlntqenvalslanigrslgmanaftrvyqlnglagyfkgigarviyqmps 246
QY 240 TAIWSVEFFKYLITKROEWRA 263
Db 247 taiswsveffkylitkrqlenra 270

RESULT 6
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX
AC AAM00938;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow protein, SEQ ID NO: 414.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX
PN WO200153453-A2.
XX
PD 26-JUL-2001.
XX
PF 23-DEC-2000; 2000WO-US34960.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
DR WPI; 2001-488707/53.
DR N-PSDB; AAH90057.
XX
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 10; Page 504-505; 648pp; English.
XX
CC The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 366 AA;

Query Match 71.1%; Score 999.5; DB 22; Length 366;
Best Local Similarity 68.6%; Pred. No. 3.1e-106;
Matches 181; Conservative 40; Mismatches 42; Indels 1; Gaps 1;

QY 1 MQSLQDPAARVNNVLEALWRIETEGLRPMRGLNVTATGAGPAHALYFACYELKKTLL 60
Db 101 mqslspdpkaqytslygalklmrtegrprlgrvnmimgagpahamyfacyenmkrtl 160

QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMPVAVKQRMQNSPYHRVTCVRAVWONE 120
Db 161 ndvfhhqgnshlangjagsmatllhdavmnpaevvqkrlqmqnsqrsaiscirtwrt 220

QY 121 GAGAFYRSYTTQLTWNVPQAIHFMTYEFQLQEHFNPQRYPNPSHVLSCACAGAAAAAT 180
Db 221 glgafyrtynpqtlmnpifqsthfityeflqeqvnphtynpqshlsgglagalaat 280

QY 181 TPLDVCKTLNTQESALN-SHITGHITGMSAFRTYQVGTAYFRGQARVYQIPS 239
Db 281 tpldvcktlntqenvalnsrlsgmanaftrvqlnglagfykgldgarvlyqmps 340

QY 240 TAIAWSVEFFKYLITTKRQEWRA 263
Db 341 taiswsveyeffkyltkqlenra 364

RESULT 7
AAU01989
ID AAU01989 standard; Protein; 174 AA.
XX
AC AAU01989;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human secreted protein encoded by gene #25.
XX
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;

KW nervous system disorder; bacterial infection; viral infection;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX
OS Homo sapiens.
XX WO200123598-AL.
XX
XX
PD 05-APR-2001.
XX
PF 26-SEP-2000; 2000WO-US26324.
XX
PR 27-SEP-1999; 99US-0155807.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis G, Ruben SM, Rosen CA;
XX
XX WPI; 2001-281684/29.
DR N-PSDB; AAS03906.
XX
PT Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Disclosure; Page 501; 518pp; English.
XX
CC The sequence represents a human secreted protein encoded by a nucleic
CC acid of the invention. Secreted proteins and their related nucleic acids
CC can be used in the diagnosis of or susceptibility to a pathological
CC condition by determining the presence or absence of a mutation in a
CC nucleic acid or the presence or amount of expression of a secreted
CC protein. The sequences are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The antibodies to the polypeptides can also be used in
CC alleviating symptoms associated with disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent
CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The peptides can also
CC be used to aid wound healing and epithelial cell proliferation, to help
CC prevent skin ageing due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues, in chemotaxis and as a food
CC additive or preservative to alter storage capabilities.
XX
SQ Sequence 174 AA;

Query Match 65.1%; Score 916; DB 22; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.3e-97;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 TLLHDAAMPVAVKQRMQNSPYHRVTCVRAVWQNEGAGAFYRSYTTQLTMTNVPQA 141
Db 1 tllhdaampvavkqrmqnsyphrvtdcvravwqnegagafysyttqltmnpvqqa 60

QY 142 IHFMTYEFQLQEHFNPQRYPNPSHVLSCACAGAAAAATPLDVCKTLNTQESALNSH 201
Db 61 ihfmyeflqehfnpqrrypnpsshvlsqacagavaaaattpldvcktlntqesalns 120

QY 202 ITGHITGMSAFRTYQVGTAYFRGQARVYQIPSTAIWSVVEFFKYLIT 255
Db 121 itghitgmsaftrvyqvgvtayfrgvarviyqipstaiwsvveyeffkylit 174

RESULT 8
AAB40634
ID AAB40634 standard; Protein; 176 AA.
XX
AC AAB40634;

XX 08-FEB-2001 (first entry)
 XX Human ORFX ORF398 polypeptide sequence SEQ ID NO:796.
 DE
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antichyroid;
 KW antanaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 XX WO200058473-A2.
 PN
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 XX Shinkets RA, Leach M;
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC74843.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 854; 5507pp; English.
 XX
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antichyroid; and antanaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. NO. 2.2e-96;

Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 VVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIFHMTYEFLOEH 153
 DB 5 VVKRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIFHMTYEFLOEH 64
 QY 154 FNPQRYNPSHVLGACAGAVAAATTPLDVCKTLTNTQESALNSHITGHTGMSAF 213
 DB 65 FNPQRYNPSHVLGACAGAVAAATTPLDVCKTLTNTQESALNSHITGHTGMSAF 124
 QY 214 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 265
 DB 125 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 176
 RESULT 9
 AAB98221
 ID AAB98221 standard; Protein; 176 AA.
 XX
 AC AAB98221;
 XX
 XX 20-AUG-2001 (first entry)
 DT
 XX Human mitochondrial solute carrier (hMSC-homologue) protein SEQ:7.
 DE
 XX Human; hMSC; hMSC-homologue; mitochondrial solute carrier.
 KW
 XX Homo sapiens.
 OS
 XX CN1281895-A.
 PN
 XX 31-JAN-2001.
 PD
 XX 27-JUN-2000; 2000CN-0116795.
 PF
 XX 27-JUN-2000; 2000CN-0116795.
 PR
 XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
 PA
 XX Li N, Xu X, Xiao H;
 PI
 XX WPI; 2001-282781/30.
 DR
 DR N-PSDB; AAB22162.
 XX
 XX New human mitochondrial solute carrier (hMSC) protein isomer and its
 PT code sequence -
 PT
 XX Claim 4; Page 18 (disclosure); 22pp; Chinese.
 PS
 XX The present invention describes a human mitochondrial solute carrier
 CC protein, designated hMSC-homologue, which is expressed in human
 CC hypophysis. Also described are methods for the preparation and
 CC detection of hMSC-homologue protein and nucleotide sequences. The
 CC present sequence represents hMSC-homologue, as given in the present
 CC invention.
 XX
 XX Sequence 176 AA;
 SQ

Query Match 64.7%; Score 910; DB 22; Length 176;
 Best Local Similarity 100.0%; Pred. NO. 2.2e-96;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 VVKORMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIFHMTYEFLOEH 153
 DB 5 VVKRMQYNPSYHRVTDCVRAVWQNEGAGAFYRSYTTQLTMTNVPFOAIFHMTYEFLOEH 64
 QY 154 FNPQRYNPSHVLGACAGAVAAATTPLDVCKTLTNTQESALNSHITGHTGMSAF 213
 DB 65 FNPQRYNPSHVLGACAGAVAAATTPLDVCKTLTNTQESALNSHITGHTGMSAF 124
 QY 214 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 265
 DB 125 RTVYQGGVTAYFRGVOARVIYQIPSTAIAWSYEFFKYLITKROEWRAGK 176
 Query Match 64.7%; Score 910; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. NO. 2.2e-96;

Db 125 rtyvggvtayfrvgarvlyqipstaiaawsyeffkylithkrqewragk 176

RESULT 10

ARB60658
ID AAB60658 standard; Protein; 331 AA.

XX
AC AAB60658;

XX
DT 04-MAY-2001 (first entry)

XX
DE Human mitochondrial solute carrier protein hMSC-o.

XX
KW Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
preparation; detection.

XX
OS Homo sapiens.

XX
PN CN1269409-A.

XX
PD 11-OCT-2000.

XX
PF 17-MAR-2000; 2000CN-0114958.

XX
PR 17-MAR-2000; 2000CN-0114958.

XX
PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX
PI Zhang X, Gao X, Xiao H;

XX
DR WPI; 2001-050544/07.

XX
DR N-PSDB; AAF59920.

XX
PT New human mitochondrion solute carrier protein and its nucleic acid -

XX
PS Claim 4; Page 20; 21pp; Chinese.

XX
CC The invention relates to a novel human mitochondrial carrier
protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
expressed in normal human hypothalamus tissue. The invention also relates
to the preparation of hMSC-o proteins and nucleic acids, and the
detection of hMSC-o proteins and nucleic acids in a sample. The present
sequence represents hMSC-o.

XX
SQ Sequence 331 AA;

Query Match 60.3%; Score 848.5; DB 22; Length 331;
Best Local Similarity 68.4%; Pred. No. 6.7e-89;
Matches 154; Conservative 34; Mismatches 36; Indels 1; Gaps 1;

QY 1 MQSLQDDPAARYRNVEALWRIITTEGLWRPRLGNVTATGAGPAHALYFACYEKLKTL 60

Db 57 mqlslpdpkaqtsiygalkimrtgefwrlgrgvnmimgagpahamyfacyenmkrtl 116

QY 61 SDVHPGGNSHANGAACGVATLLHDAAMNPAEVVKORMQYNPSVHRVDCVRAYWONE 120

Db 117 ndvfhhqgshlanglagsmatllhdavmpaevvkqrlqmynsqhrsaiscirtvwrt 176

QY 121 GAGAFYRSYTTQLTNMVFPQAIHFMTYEFLOEHFNPRRYNPSHVLGACAGAVAAAT 180

Db 177 glgafyrsyttqtmnlpfqsibhfyeflqeqvnpbrtynpqshisgglagalaat 236

QY 181 TPLDVCKTLNLTQESLALN-SHITGHITCMASAFRTVYQVGWTA 224

Db 237 tpldvcktlntqenvalslanlsgrlsmanaftrvyqlnglpa 281

RESULT 11

AAU29748

ID AAU29748 standard; Protein; 677 AA.

XX
AC AAU29748;

XX

DT 18-DEC-2001 (first entry)

XX
DE Novel human secreted protein #239.

XX
KW Human; vaccination; gene therapy; nutritional supplement;
stem cell proliferation; haematopoiesis; nerve tissue regeneration;
immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX
OS Homo sapiens.

XX
PN WO200179449-A2.

XX
PD 25-OCT-2001.

XX
PF 16-APR-2001; 2001WO-US08656.

XX
PR 18-APR-2000; 2000US-0552929.

XX
PR 26-JAN-2001; 2001US-0770160.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Liu C, Drmanac RT;

XX
DR WPI; 2001-611725/70.

XX
CC Nucleic acids encoding a range of human polypeptides, useful in genetic
vaccination, testing and therapy -

XX
PS Claim 20; Page 189; 765pp; English.

XX
CC The invention relates to novel human secreted polypeptides. The
polypeptides and antibodies to the polypeptides are useful for
determining the presence of or predisposition to a disease associated
with altered levels of polypeptide. The polypeptides are also useful for
identifying agents (agonists and antagonists) that bind to them. Cells
expressing the proteins are useful for identifying a therapeutic agent
for use in treatment of a pathology related to aberrant expression or
physiological interactions of the polypeptide. Vectors comprising
the nucleic acids encoding the polypeptides and cells genetically
engineered to express them are also useful for producing the proteins.
The proteins are useful in genetic vaccination, testing and
therapy, and can be used as nutritional supplements. They may be used to
increase stem cell proliferation; to regulate haematopoiesis; and in
bone, cartilage, tendon and/or nerve tissue growth or regeneration;
immune suppression and/or stimulation; as anti-inflammatory agents; and
in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
sequences of novel human secreted proteins of the invention.

XX
SQ Sequence 677 AA;

Query Match 53.7%; Score 754.5; DB 22; Length 677;
Best Local Similarity 66.7%; Pred. No. 1.3e-77;
Matches 138; Conservative 30; Mismatches 38; Indels 1; Gaps 1;

QY 1 MQSLQDDPAARYRNVEALWRIITTEGLWRPRLGNVTATGAGPAHALYFACYEKLKTL 60

Db 85 mqlslpdpkaqtsiygalkimrtgefwrlgrgvnmimgagpahamyfacyenmkrtl 144

QY 61 SDVHPGGNSHANGAACGVATLLHDAAMNPAEVVKORMQYNPSVHRVDCVRAYWONE 120

Db 145 ndvfhhqgshlanglagsmatllhdavmpaevvkqrlqmynsqhrsaiscirtvwrt 204

QY 121 GAGAFYRSYTTQLTNMVFPQAIHFMTYEFLOEHFNPRRYNPSHVLGACAGAVAAAT 180

Db 205 glgafyrsyttqtmnlpfqsibhfyeflqeqvnpbrtynpqshisgglagalaat 264

QY 181 TPLDVCKTLNLTQESLALN-SHITGHI 206

Db 265 tpldvcktlntqenvalslanlsgrl 291

```
RESULT 12
AAM80023
ID AAM80023 standard; Protein; 199 AA.
XX
AC AAM80023;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3669.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Wang Y, Wejhrman T, Goodrich R;
XX
WPI; 2001-476283/51.
DR N-PSDB; AAK53156.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
Claim 20; Page 409-410; 6221pp; English.
XX
The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 199 AA;
```

```
Query Match 52.68; Score 739.5; DB 22; Length 199;
Best Local Similarity 71.68; Pred. No. 1.2e-76;
Matches 136; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 75 GAAGCVATLLHDAAMNPAEVVKORMQNSPYHRVTDCVRAVMQNEGAFYRSYTTQLT 134
   ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 8 gagsmatllhdavmnpaevvkqrlqmynsqhrsaiscirtvrtglgafyrsyttqlt 67
QY 135 MNVPFQAIHFMTYBFLQEHFNQRRYNPSSHVLSGACAGAVAAAATPLDVCCKTLNTQE 194
```

```
Db 68 mnipfqsinfityeflqeqvnphtynpqshisgslagalaataatpldvcktlIntge 127
QY 195 SLALN-SHITGHTCMASAFRTVYQVGVTAYFRCVQARVYQIPSTAIANSVYEFKYL 253
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 128 nvalslanisgrlsgrlmanaftrvqqlnglagfykgigqarvlyqmpstaisvyeffkyf 187
QY 254 ITRKQEWRA 263
   :|||| |||
Db 188 ltkrqlenra 197

RESULT 13
AAM93909
ID AAM93909 standard; Protein; 187 AA.
XX
AC AAM93909;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 4059.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EPI130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0193765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
WPI; 2001-524255/58.
DR N-PSDB; AAK94871.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
Claim 8; SEQ ID NO 4059; 1380pp + sequence listing; English.
XX
The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 187 AA;
```

```
Query Match 51.68; Score 725.5; DB 22; Length 187;
Best Local Similarity 71.98; Pred. No. 4.3e-75;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;
```

```
QY 80 VATTLLHDAAMNPAEVVKORMQNSPYHRVTDCVRAVMQNEGAFYRSYTTQLTMNVPF 139
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 matllhdavmnpaevvkqrlqmynsqhrsaiscirtvrtglgafyrsyttqtmnpf 60
```

Qy	140	QAIHFMTYEFLOEHFNPORRYNPPSSHVLSGACAGAVAAAAATTPLDVCKTLLNTQESIALN	199
Db	61	qsihfiyeflqeqvnpshrtnpqshisgglagalaataatpldvcktkllntgenvals	120
Qy	200	-SHITGHTGMASAFRTVYQVGGVTAYFRGVQVARVIQIPSTAIASVSYEFFKYLITKQ	258
Db	121	lanisgrlsgmanaftrvqnglqagykqigarviqmpstaisvsvyeffkyfktkrq	180
Qy	259	EEWRA 263	
Db	181	lenra 185	
RESULT 14			
AA	98222		
ID	AA98222	standard; Protein; 187 AA.	
XX	AC	AA98222;	
XX	DT	20-AUG-2001 (first entry)	
XX	DE	Human mitochondrial solute carrier (hMSC) protein sequence.	
XX	KW	Human; hMSC; hMSC-homologue; mitochondrial solute carrier.	
XX	OS	Homo sapiens.	
XX	PN	CN1281895-A.	
XX	PD	31-JAN-2001.	
XX	PF	27-JUN-2000; 2000CN-0116795.	
XX	PR	27-JUN-2000; 2000CN-0116795.	
XX	PA	(NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.	
XX	Li N, Xu X, Xiao H;		
XX	WPI;	2001-282781/30.	
XX	DR		
XX	PT	New human mitochondrial solute carrier (hMSC) protein isomer and its	
XX	code sequence -		
XX	PS	Example 2; Fig 3; 22pp; Chinese.	
XX	CC	The present invention describes a human mitochondrial solute carrier	
XX	CC	protein, designated hMSC-homologue, which is expressed in human	
XX	CC	hypophysis. Also described are methods for the preparation and	
XX	CC	detection of hMSC-homologue protein and nucleotide sequences. The	
XX	CC	present sequence represents the human mitochondrial solute carrier	
XX	CC	hMSC protein sequence which is given in comparison with the	
XX	CC	hMSC-homologue protein sequence in an example from the present	
XX	CC	invention.	
XX	SQ	Sequence 187 AA;	
Query Match 51.6%; Score 725.5; DB 22; Length 187;			
Best Local Similarity 71.9%; Pred. No. 4.3e-75;			
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps			
Qy	80	VATLLHDAANPAEVVKORMQMYNPSVHRVTCVRAVMQNEGAGAFYRSTTOLTWNVPF	139
Db	1	matllhdavmnpaevvkvrlqmynsqhrsaiscirtvrtreglgafrysttqltmnlpf	60
Qy	140	QAIHFMTYEFLOEHFNPORRYNPPSSHVLSGACAGAVAAAAATTPLDVCKTLLNTQESIALN	199
Db	61	qsihfiyeflqeqvnpshrtnpqshisgglagalaataatpldvcktkllntgenvals	120
Qy	200	-SHITGHTGMASAFRTVYQVGGVTAYFRGVQVARVIQIPSTAIASVSYEFFKYLITKQ	258
Db	121	lanisgrlsgmanaftrvqnglqagykqigarviqmpstaisvsvyeffkyfktkrq	180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:50 ; Search time 48.11 Seconds
(without alignments)
134.541 Million cell updates/sec

Title: US-09-870-113-8

Perfect score: 1406

Sequence: 1 MQSLQDPDPAARYNVLEALW.....VYEFKYLITKROENRAGK 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	429.5	30.5	289	4	US-09-068-140A-10
2	429.5	30.5	328	4	US-09-068-140A-15
3	251.5	17.9	351	2	US-08-933-750C-19
4	251.5	17.9	351	4	US-09-234-613-19
5	219.5	15.6	469	4	US-09-188-930-339
6	213	15.1	447	4	US-09-160-119-4
7	213	15.1	674	4	US-09-160-119-2
8	212.5	15.1	320	2	US-08-933-750C-12
9	212.5	15.1	320	4	US-09-234-613-12
10	168.5	12.0	432	2	US-08-937-466-4
11	168.5	12.0	432	2	US-09-172-528-4
12	168.5	12.0	432	3	US-09-318-199-4
13	168.5	12.0	432	4	US-09-503-579-4
14	168	11.9	311	2	US-08-775-009-32
15	166.5	11.8	311	2	US-08-775-009-33
16	164.5	11.7	308	2	US-08-937-466-2
17	164.5	11.7	308	2	US-09-172-528-2
18	164.5	11.7	308	3	US-09-318-199-2
19	164.5	11.7	308	4	US-09-503-579-2
20	163.5	11.6	306	5	PCT-US94-09799-1
21	161	11.5	312	4	US-09-142-565-2
22	154	11.0	307	2	US-08-807-861A-56
23	154	11.0	307	3	US-09-210-681-56
24	154	11.0	307	3	US-08-946-719A-56
25	148	10.5	256	2	US-08-937-466-6
26	148	10.5	256	2	US-09-172-528-6
27	148	10.5	256	3	US-09-318-199-6

28 148 10.5 256 4 US-09-503-579-6 Sequence 6, Appli
29 137 9.7 299 1 US-08-518-878B-56 Sequence 56, Appl
30 137 9.7 299 2 US-08-470-868A-56 Sequence 56, Appl
31 137 9.7 309 1 US-08-518-878B-51 Sequence 51, Appl
32 137 9.7 309 2 US-08-807-861A-51 Sequence 51, Appl
33 137 9.7 309 2 US-08-470-868A-51 Sequence 51, Appl
34 137 9.7 309 3 US-09-210-681-51 Sequence 51, Appl
35 137 9.7 309 3 US-08-946-719A-51 Sequence 51, Appl
36 121.5 8.6 298 3 US-08-961-871-10 Sequence 10, Appl
37 121.5 8.6 312 4 US-09-188-930-142 Sequence 142, App
38 106 7.5 303 1 US-08-518-878B-37 Sequence 37, Appl
39 106 7.5 303 1 US-08-294-522B-36 Sequence 36, Appl
40 106 7.5 303 2 US-08-807-861A-37 Sequence 37, Appl
41 106 7.5 303 2 US-08-470-868A-37 Sequence 37, Appl
42 106 7.5 303 3 US-09-210-681-37 Sequence 37, Appl
43 106 7.5 303 3 US-08-946-719A-37 Sequence 37, Appl
44 85 6.0 1832 3 US-09-335-409-4 Sequence 4, Appli
45 85 6.0 1832 4 US-09-568-102-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-068-140A-10
; Sequence 10, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinher, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-10

```
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match 30.5%; Score 429.5; DB 4; Length 289;
Best Local Similarity 38.7%; Pred. No. 1.6e-42;
Matches 99; Conservative 38; Mismatches 108; Indels 11; Gaps 4;

QY 9 AARYRNVLALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDVTHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 NSHIANGAAGCVATLLHDAAMNPAEYVVKORMQNSPYHRVTDVCRVAVQNEGAGAFYRS 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNGAHAVSGVFATVADAVITPMDVVKQRLQSSPYKGVVDVCRVRLVEEGIGAFYAS 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 YTTQLTMNVPFOAIHPTWTFEFLQE---HFNPO---RRYNPSSHVLSGACAGAVAAAATTP 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 YRTTVVMNAPFTAVHPTATYATKKGLEVPETANDENLLVHATAGAAALAAVVTTP 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 DVCKTLLNTQESLALNSHTGHTGMASAFRTVYQVGGVTATFRGVOARVYIOPSTATA 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 DVVKTLQCCQGVCGCDRFSSSSIQDVIG---SIVKNGYVGLMRGWIPLMLFHAPAAAI 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 WSVYEFKYLITKROE 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 WSTYEASKTFFOKLNE 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: NO. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match 30.5%; Score 429.5; DB 4; Length 328;
Best Local Similarity 38.7%; Pred. No. 2e-42;
Matches 99; Conservative 38; Mismatches 108; Indels 11; Gaps 4;

QY 9 AARYRNVLALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDVTHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 SAQSAGLRQALGSILKVEGPAGLYRGIGAMGLGAGPAHAHVFSVYEMCKETFS---HGDP 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 NSHIANGAAGCVATLLHDAAMNPAEYVVKORMQNSPYHRVTDVCRVAVQNEGAGAFYRS 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 SNGAHAVSGVFATVADAVITPMDVVKQRLQSSPYKGVVDVCRVRLVEEGIGAFYAS 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 YTTQLTMNVPFOAIHPTWTFEFLQE---HFNPO---RRYNPSSHVLSGACAGAVAAAATTP 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 YRTTVVMNAPFTAVHPTATYATKKGLEVPETANDENLLVHATAGAAALAAVVTTP 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 DVCKTLLNTQESLALNSHTGHTGMASAFRTVYQVGGVTATFRGVOARVYIOPSTATA 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 DVVKTLQCCQGVCGCDRFSSSSIQDVIG---SIVKNGYVGLMRGWIPLMLFHAPAAAI 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 WSVYEFKYLITKROE 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 WSTYEASKTFFOKLNE 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-08-933-750C-19
; Sequence 19, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lai, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536.
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
```


Db 218 ASRSNNNCIVGGFTQMTIREGAKSLWRG-NGINVLK--IAPESAIKPMAYEQMKRLVGS 274
QY 63 V-----IHPGNSHTANGAAGCVATLLHDAAMNPAEVVKORMQVNS-PYHRVTDVCVRAVW 117
Db 275 QETLRIH---ERLVAGSLAGAIA-----QSSLYPMEVLKTRMALNKTQYSGMLDCARRIL 327
QY 118 QNEGAGAFYSYTTQLTMTNVPFOAIHMTYEFLEQEHFNPPORRY-----NPSSSHVL--SGA 170
Db 328 AKEGVAAYKGYIPNMLGIIPYAGIDLAVYETLKNTW--LQRAVNSADPGVFLACGT 385
QY 171 CAGAVAAATTPDVCKTLNTQESLALNSHITGTMASAFRTVYQVGGVTAYFRGVQ 230
Db 386 ISSTCGLASYPLALVTRMQAASIEGAPEVT-----MSSLFKQILRTGAGFLYRGLA 440
QY 231 ARVIQIPSTAIANSVVEFFK 251
Db 441 PFMKVIPAVSISVYVENLK 461

RESULT 6
US-09-160-119-4
; Sequence 4, Application US/09160119A
; Patent No. 6316219
; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160,119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 447
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-4

Query Match 15.1%; Score 213; DB 4; Length 447;
Best Local Similarity 26.7%; Pred. No. 1.4e-16;
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS DVI-----HP 66
Db 143 YKNSFDCFKKVLRYEGFGLYRGLLPOLLGVAPEK-----IKLTVNDFVRDKFMHK 194
QY 67 GGN-----SHIANGAAGCVATLLHDAAMNPAEVVKORMQVNSPYHRVTDVCR-----AVW 117
Db 195 DGSVPLAAEILAGCAGGSQVIF-----TNPLEIVKIRLQVAG-----EITGPRVSALS 246
QY 118 QNEGAGAFYSYTTQLTMTNVPFOAIHMTYEFLEQEHF-NPQRRYNPSSHVLSGACAGAVA 176
Db 247 RDLGFGGIYKGAKACFLRDIPFSAIYPCYAHVKASFANEDGQVSPGSLLAGAIGMPA 306
QY 177 AAATTPDVCKTLNTQESLALNSHITGTMASAFRTVYQVGGVTAYFRGVQARVIYQ 236
Db 307 ASLVTADVTKIRLQVAARAGQTTY-----SGVIDCFKILREEGPKALKWGAGARVFRS 361
QY 237 IPSTAIANSVVEFFK 251
Db 362 SPQFGVTLTYELLQ 376

RESULT 7
US-09-160-119-2
; Sequence 2, Application US/09160119A
; Patent No. 6316219

; GENERAL INFORMATION:
; APPLICANT: KRIEF, STEPHANE
; APPLICANT: SOUCHET, MICHEL
; APPLICANT: BRIL, ANTOINE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30985
; CURRENT APPLICATION NUMBER: US/09/160,119A
; CURRENT FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: EP 97402511.6
; EARLIER FILING DATE: 1997-10-23
; EARLIER APPLICATION NUMBER: EP 98401655.0
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-160-119-2

Query Match 15.1%; Score 213; DB 4; Length 674;
Best Local Similarity 26.7%; Pred. No. 2.7e-16;
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLS DVI-----HP 66
Db 370 YKNSFDCFKKVLRYEGFGLYRGLLPOLLGVAPEK-----IKLTVNDFVRDKFMHK 421
QY 67 GGN-----SHIANGAAGCVATLLHDAAMNPAEVVKORMQVNSPYHRVTDVCR-----AVW 117
Db 422 DGSVPLAAEILAGCAGGSQVIF-----TNPLEIVKIRLQVAG-----EITGPRVSALS 473
QY 118 QNEGAGAFYSYTTQLTMTNVPFOAIHMTYEFLEQEHF-NPQRRYNPSSHVLSGACAGAVA 176
Db 474 RDLGFGGIYKGAKACFLRDIPFSAIYPCYAHVKASFANEDGQVSPGSLLAGAIGMPA 533
QY 177 AAATTPDVCKTLNTQESLALNSHITGTMASAFRTVYQVGGVTAYFRGVQARVIYQ 236
Db 534 ASLVTADVTKIRLQVAARAGQTTY-----SGVIDCFKILREEGPKALKWGAGARVFRS 588
QY 237 IPSTAIANSVVEFFK 251
Db 589 SPQFGVTLTYELLQ 603

RESULT 8
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/933,750C
;; FILING DATE: September 23, 1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0356 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 320 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: SPLNOT02
;; CLONE: 207452
;; US-08-933-750C-12

Query Match 15.1%; Score 212.5; DB 2; Length 320;
Best Local Similarity 23.4%; Pred. No. 9.6e-17;
Matches 62; Conservative 54; Mismatches 116; Indels 33; Gaps 9;
QY 7 DPAARYNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDV 63
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 54 DPSAKYHGLQASRILOEGPTAFWKGHVPAQILSIGYGAQVFLSF-----EMTEL 106
QY 64 IHFGG-----NSH-IANGAAGCVATLLHDAAMNPAEVVKQRMQYNSP--YHRVTDVC 113
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 107 VHRGSVVDAREFSVHFVCGGLAACMATL---TVHPVDVLRTRFAAQGEPKVYNTLRHAV 162
QY 114 RAWQNEGAGAFRSYTTQLTMTNVPQAIHFMYEFLQEHFN-----PQRRYPSSHVL 168
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 163 GTMYRSEGPQVFKGLAPTLIAIFPYAGLQFSCYSSLKHLKWAIPAEGKKKNENLQNL 222
QY 169 GACAGAVAAATTPLDVCKTLLNT---QESIALNSHTTGHTGMASAFRTVYQVGGVTAY 225
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 223 GSGAGVISKTLTLPDLFKRLQVGGFEHARAAGQVRRY-KGLMDCAKQVLOKEGALGF 281
QY 226 FRGQARVIYQIPSTAIAMSVYEFF 250
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 282 FKGLSPSLKALSTGMFMSYEFF 306

RESULT 9
US-09-234-613-12
; Sequence 12, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/234,613
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/933,750
;; FILING DATE: September 23, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0356 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 320 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: SPLNOT02
;; CLONE: 207452
;; US-09-234-613-12
Query Match 15.1%; Score 212.5; DB 4; Length 320;
Best Local Similarity 23.4%; Pred. No. 9.6e-17;
Matches 62; Conservative 54; Mismatches 116; Indels 33; Gaps 9;
QY 7 DPAARYNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDV 63
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 54 DPSAKYHGLQASRILOEGPTAFWKGHVPAQILSIGYGAQVFLSF-----EMTEL 106
QY 64 IHFGG-----NSH-IANGAAGCVATLLHDAAMNPAEVVKQRMQYNSP--YHRVTDVC 113
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 107 VHRGSVVDAREFSVHFVCGGLAACMATL---TVHPVDVLRTRFAAQGEPKVYNTLRHAV 162
QY 114 RAWQNEGAGAFRSYTTQLTMTNVPQAIHFMYEFLQEHFN-----PQRRYPSSHVL 168
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 163 GTMYRSEGPQVFKGLAPTLIAIFPYAGLQFSCYSSLKHLKWAIPAEGKKKNENLQNL 222
QY 169 GACAGAVAAATTPLDVCKTLLNT---QESIALNSHTTGHTGMASAFRTVYQVGGVTAY 225
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 223 GSGAGVISKTLTLPDLFKRLQVGGFEHARAAGQVRRY-KGLMDCAKQVLOKEGALGF 281
QY 226 FRGQARVIYQIPSTAIAMSVYEFF 250
||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:| ||:|:|
Db 282 FKGLSPSLKALSTGMFMSYEFF 306

RESULT 10
US-08-937-466-4
; Sequence 4, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA

```
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-937-466-4
```

```
Query Match 12.0%; Score 168.5; DB 2; Length 432;
Best Local Similarity 24.6%; Pred. No. 2.5e-11;
Matches 66; Conservative 38; Mismatches 121; Indels 43; Gaps 10;

Qy 9 AARYRNVEALWRIIRTEGLRPMRGLNVATGAGPAHALYFA-----CYEKLKKTLSDV 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 SVQYRGVLGTILTWRTGSPRPSYGL-----VAGLHRQMSFASIRIGLYDSVKQFY-- 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 IHPGNSHIANG---AAGCVATLLHDAAMNPAEVVKQRMQY-----NSPYHRVTDVCV 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 --PKGADHSSVAIRLAGCTTGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKYRTMDAY 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 RAVMNEGAGAFYRSYTTQLTMNVFPQAIHFMTYEFLEQHEHNPQRRY--NPSSHVLSCAC 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 RTIAREEGVRLGKGTWPNITRNAIVNCAEMVTYDIIEKLLSHLFTDNFPCHFVSAGF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 172 AGAVAAATATPLDVCCKTLNTQESLALNSHITGHTMASAFRTVYQGGVTAYFRGVQA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 AGFCATVVASPDVVVKTRY-----MNAPLGRYSPLHCLMKWVAQ-EGPTAFYKGF-- 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 232 RVYIQIPS--TAIAWSVYEFKYLITKR 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----VPSFLRLGAWNMVMFVTEOLKR 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 11
US-09-172-528-4
; Sequence 4, Application US/09172528
; Patent No. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-172-528-4

Query Match 12.0%; Score 168.5; DB 2; Length 432;
Best Local Similarity 24.6%; Pred. No. 2.5e-11;
Matches 66; Conservative 38; Mismatches 121; Indels 43; Gaps 10;

Qy 9 AARYRNVEALWRIIRTEGLRPMRGLNVATGAGPAHALYFA-----CYEKLKKTLSDV 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 SVQYRGVLGTILTWRTGSPRPSYGL-----VAGLHRQMSFASIRIGLYDSVKQFY-- 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 IHPGNSHIANG---AAGCVATLLHDAAMNPAEVVKQRMQY-----NSPYHRVTDVCV 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 --PKGADHSSVAIRLAGCTTGAMAVTCAQPTDVVVKVRFQAMIRLGTGGERKYRTMDAY 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 RAVMNEGAGAFYRSYTTQLTMNVFPQAIHFMTYEFLEQHEHNPQRRY--NPSSHVLSCAC 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 RTIAREEGVRLGKGTWPNITRNAIVNCAEMVTYDIIEKLLSHLFTDNFPCHFVSAGF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 172 AGAVAAATATPLDVCCKTLNTQESLALNSHITGHTMASAFRTVYQGGVTAYFRGVQA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 AGFCATVVASPDVVVKTRY-----MNAPLGRYSPLHCLMKWVAQ-EGPTAFYKGF-- 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 232 RVYIQIPS--TAIAWSVYEFKYLITKR 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----VPSFLRLGAWNMVMFVTEOLKR 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 12
US-09-318-199-4
; Sequence 4, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```


; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-775-009-32

Query Match 11.9%; Score 168; DB 2; Length 311;
Best Local Similarity 24.2%; Pred. No. 1.7e-11;
Matches 65; Conservative 38; Mismatches 118; Indels 48; Gaps 8;
QY 11 RYRVLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTLSDVIHPGNGS 70
DB 63 RYRGIGDCVQRTVRSHGLYRGLSLYLSIPKAAVRFGMFEFLSNHMDA---QGRLL 119
QY 71 HIANG-----AAGCVATLLHDAAMNPAEVVVKQM-----QMYNSPYHRTDCVRVW 117
DB 120 DSTRGLLGLGAGVABV---VCPMTVKVKFIHDTSPNPKYRGFFHGVREIYR--- 173
QY 118 QNEGAGAFYSYTTQLTMTNVPFOAIHFTMYEFLQEHF---NFORRYNPSHVLGACAGA 174
DB 174 -EGLKGTGYOGLTATVLKOGSNOAIFRFVMTSLRNWYRGDNPNKPNPLITGVFGAIGA 232
QY 175 VAAATTPLDVCKTLTNTQESLALNSHITGHTGMASAFRTVYQVG-----GVTAIYR 227
DB 233 ASVFGNTPLDVIKTRMOGLE-----AHKYRNTWDCGLQILKKEGLKAFYK 277
QY 228 GVOARVIYQIPSTAIWSVYEFKYLITK 256
DB 278 GTFPRGLGRVCLDVAIVFYDEVYKLLNK 306

RESULT 15
US-08-775-009-33
; Sequence 33, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Wellong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Velio cardiolfacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-775-009-33

Query Match 11.8%; Score 166.5; DB 2; Length 311;
Best Local Similarity 24.7%; Pred. No. 2.6e-11;
Matches 69; Conservative 35; Mismatches 124; Indels 51; Gaps 8;
QY 4 LQDPDPA---RYRVLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLLKTL 60
DB 53 LQIDERANPPRYRGIGDCVQRTVRSHGLYRGLSLYLSIPKAAVRFGMFEFLSNHM 112
QY 61 SDV-----IHPGNSHIANGAACCVATLLHDAAMNPAEVVVKQM-----QMYNSPYH 107
DB 113 RDAQRLDSRGLLGLGAGVABV---PMTVKVKFIHDTSSNPKNKRGFFH 166
QY 108 RYTDVYRVWQNEGAGAFYSYTTQLTMTNVPFOAIHFTMYEFLQEHF---NFORRYNPS 164
DB 167 GVREIYR---EGLKGTGYOGLTATVLKOGSNOAIFRFVMTSLRNWYQDNPNKPNPLI 222
QY 165 HVLGACAGAVAAATTPLDVCKTLTNTQESLALNSHITGHTGMASAFRTVYQVG----- 220
DB 223 TGVEGAVAGAASVFGNTPLDVIKTRMOGLE-----AHKYRNTLDCGVQIL 267
QY 221 ---GVTAIYRFGVQARVIYQIPSTAIWSVYEFKYLITK 256
DB 268 KNEGPKAFYKGVPRGLGRVCLDVAIVFYDEVYKLLNK 306

Search completed: August 27, 2002, 03:58:51
Job time: 4893 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:57:41 ; Search time 80.04 Seconds
(without alignments)
318.137 Million cell updates/sec

Title: US-09-870-113-8
Perfect score: 1406
Sequence: 1 MQSLQDPDAARYNVLEALW.....VVEFFKYLITKQEEWRAGK 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	677	48.2	312	2 T26089	hypothetical prote
2	467	33.2	303	2 T39149	probable RNA splic
3	441	31.4	304	2 S13533	mtRNA splice defec
4	431.5	30.7	781	2 A86205	hypothetical prote
5	428	30.4	314	2 S55179	mtRNA splice defec
6	414	29.4	331	2 T00582	probable mitochond
7	283	20.1	328	2 T19322	hypothetical prote
8	275.5	19.6	368	2 S54524	hypothetical prote
9	251.5	17.9	351	2 T43493	hypothetical prote
10	248.5	17.7	367	2 G96770	hypothetical prote
11	242.5	17.2	394	1 S44092	probable carrier p
12	226	16.1	284	2 S45458	PET8 protein - yea
13	224.5	16.0	902	1 S54495	probable carrier p
14	219.5	15.6	269	2 T02990	hypothetical prote
15	216.5	15.4	330	2 T09362	hypothetical prote
16	213	15.1	336	2 T01839	hypothetical prote
17	210	14.9	352	2 T01729	mitochondrial solu
18	209.5	14.9	436	2 J01459	Bt1 protein precur
19	209.5	14.9	475	2 T50866	peroxisomal Ca-dep
20	208.5	14.8	392	2 T05350	adenylate transloc
21	208	14.8	366	2 S64589	hypothetical prote
22	206	14.7	322	2 S57116	probable carrier p
23	204	14.5	271	2 S62485	probable mitochond
24	203.5	14.5	338	2 T40968	probable mitochond
25	201.5	14.3	349	2 B96753	hypothetical prote
26	200	14.2	479	2 T49871	peroxisomal Ca-dep
27	199.5	14.2	307	2 S60949	probable phosphate
28	198	14.1	345	2 T37576	probable mitochond
29	196.5	14.0	326	2 S57544	probable membrane

30 196 13.9 358 2 D84901
31 193 13.7 358 2 T45934
32 192 13.7 335 2 T50393
33 191 13.6 588 2 T22688
34 188 13.4 288 2 T51595
35 188 13.4 309 2 T48156
36 187.5 13.3 296 2 B96830
37 186 13.2 332 2 T47703
38 186 13.2 373 2 S48451
39 185 13.2 325 2 T04273
40 184 13.1 300 2 S50556
41 184 13.1 335 2 S50453
42 181 12.9 500 2 T39385
43 179.5 12.8 650 2 T32897
44 179 12.7 258 2 T39347
45 178 12.7 277 2 T40033

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C;Accession: T26089

R;Swinburne, J.; Ahnscough, R.

submitted to the EMBL Data Library, October 1995

A;Reference number: Z20149

A;Accession: T26089

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-312 <WIL>

A;Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9

A;Experimental source: clone W02B12

C;Genetics:

A;Gene: CESP:W02B12.9

A;Map position: 2

A;Introns: 18/3; 251/3; 286/3

C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 48.2%; Score 677; DB 2; Length 312;

Best Local Similarity 50.6%; Pred. No. 7.7e-54;

Matches 133; Conservative 47; Mismatches 73; Indels 10; Gaps 3;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWRPMLNVTATGAGPAHALYFACYEKKKTL 60

Db 45 MQSLCPCEPKCPVHSLMSIVKREGWRLRGVNAVAAGSMFAHALYFYVEKMGYL 104

QY 61 SDVIHPGNGSH-IANGAAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVTDVCYRAYWQN 119

Db 105 TG--NSAGHSNTLAYGASGVVATLIHDAIMNPAEVVKORMQAFSPYSSLSLCARCVNR 162

QY 120 EGAGAFYRSYTTQITMNVPPQAHFYFYFLEQHFNPQRNPNPSSHVLSCACAGAVAAA 179

Db 163 EGVAIFYRSYTTQIAMNVPPQAHFYFYFLEQHFNPQRNPNPSSHVLSCACAGAVAAA 222

QY 180 TPLDVCCKTLLNTQESLANS-----HITGHITGMASAFRTVYQVGGVTAYFRGVQAR 232

Db 223 TTPMDCVKTVLNTQQAADPANRIFLQARYRYRGISDAVRTIYSORGLSGFCGLQAR 282

QY 233 VIVQIPSTAIWSYEFYFKYLIT 255

Db 283 VIFQVPATALSWSYELFKFMLS 305

RESULT 2

T39149

probable RNA splicing protein/mitochondial carrier protein - fission yeast (Schizosacc

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T39149
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21748
A:Accession: T39149
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>
A:Cross-references: EMBL:Z99168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A:Experimental source: strain 972h-; cosmid c8c9
C:Genetics:
A:Gene: SPDB:SPAC8C9.12c
A:Map position: 1
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 33.2%; Score 467; DB 2; Length 303;
Best Local Similarity 37.0%; Pred. No. 8.9e-35;
Matches 90; Conservative 55; Mismatches 92; Indels 6; Gaps 2;
QY 14 NVLEALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKLTSLDSVHPGGNSHTA 73
DB 60 NIVNSVIKISSIEGVYSLWRGSISSVMGAPSHAIYFSVLEFFSK-----INASPDRLPA 115
QY 74 NGAACCVATLLHDAAMNPAEVVKQRMQYNSPYHRVTDCAVQWQNEGAGAFYRSYTTQL 133
DB 116 SALAGACAITISDAFTPPDVVKQRMQLPSRKYKSALHCATTVERNEGIGAFYISYPTCI 175
QY 134 TNWVPQAIHFMTYEFELQEHFNQRRYNSPSSVLSGACAGAAVAAATPLDVCCKILLNQ 193
DB 176 AMSIPTAIQVATYDTCMSFLNPNAVYDPTSHISGGLSGAIASSLTPLDVCCKILLNQ 235
QY 194 ESLANSHITGHTMASAFRTVYQVGGVTAYFRGVQARVYQIIPSTATAWSVVEFFKYL 253
DB 236 GSSSTPE--VRKCKSLDVRVIYNGGIPSFYKIRPRWVMPATAVSWAAYEAGKEI 293
QY 254 ITK 256
DB 294 LIR 296

RESULT 3
S13533
mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR052c
C:Species: Saccharomyces cerevisiae
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
C:Accession: S13533; S38126
R:Wiesenberg, G.; Link, T.A.; von Ahsen, U.; Waldbherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A:Title: MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the ADP/ATP carrier family
A:Reference number: S13532; MUID:91108815
A:Accession: S13533
A:Molecule type: DNA
A:Residues: 1-304 <JMO>
A:Cross-references: EMBL:X56444; NID:g3995; PIDN:CRA39828.1; PID:g3996
R:Visiers, S.; Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38118
A:Accession: S38126
A:Molecule type: DNA
A:Residues: 1-304 <VIS>
A:Cross-references: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MRS4
A:Cross-references: SGD:S0001760; MIPS:YKR052c
A:Map position: 11R

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:20-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 31.4%; Score 441; DB 2; Length 304;
Best Local Similarity 38.4%; Pred. No. 2.1e-32;
Matches 93; Conservative 41; Mismatches 106; Indels 2; Gaps 1;
QY 15 VLEALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKLTSLDSVHPGGNSHTAN 74
DB 64 MISQISKISTMEGSMALMGVOSVILGAGPAHAHVFGTYEFCFKARLISPEDMQTHQPMKT 123
QY 75 GAAGCVATILLHDAAMNPAEVVKQRMQYNSPYHRVTDCAVQWQNEGAGAFYRSYTTQLT 134
DB 124 ALSGIIATIAADALNPDFTVKQRLQLDN--LRVNVNVTQIYQNEGFAAFYSYPTTLA 181
QY 135 MNVPQAIHFMTYEFELQEHFNQRRYNSPSSVLSGACAGAAVAAATPLDVCCKILLNQ 194
DB 182 MNIPFAAFNFIYESASAKFFNQSNPLIHCLCGGISGATCAALTPLDCKITVLQVRG 241
QY 195 SLANSHITGHTMASAFRTVYQVGGVTAYFRGVQARVYQIIPSTATAWSVVEFFKYL 254
DB 242 SETVSIEIMKDANTEGRASRAILEVHGKGFWRGLKPRIVANIPATAISWTAYEACAKHEL 301
QY 255 TK 256
DB 302 MK 303

RESULT 4
A86205
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86205
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maili, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86205
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-781 <STO>
A:Cross-references: GB:AB005172; NID:g8954043; PIDN:AAF82217.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 30.7%; Score 431.5; DB 2; Length 781;
Best Local Similarity 40.0%; Pred. No. 5e-31;
Matches 104; Conservative 35; Mismatches 104; Indels 17; Gaps 6;
QY 1 MQSLOPDPAARYNVLEALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKLT 60
DB 64 MQALRCP- LKPVGIREAFRSIIQKEGFSALYRGITWAMGLGAGPAHAHVYFSFYSKYL 122
QY 61 SDVIHPGG--NSHTANGAGCVATILLHDAAMNPAEVVKQRMQYNSPYHRVTDCAVQW 118
DB 123 S-----ACQNNVAHAMSGVGFATISSDAVTFPMDNVKQRLQMGEGTYKGVWDCKVRVLR 177
QY 119 NEGAGAFYRSYTTQLTMNVPPQAIHFMTYEFELQEHFNQRRYNSPSSVLSGACAGAAVAAATPLDVCCKILLNQ 172
DB 178 EEGIGAFYASRYTTVMNAPTAHVHATYEAARKGLMEFSPDRISDESGWLHVHATAGAA 237
QY 173 GAVAAAATPLDVCCKILLNTQESIALNSHITGHTMASAFRTVYQVGGVTAYFRGVQAR 232
DB 238 GGLAAAVTPLDVCCKILLNTQESIALNSHITGHTMASAFRTVYQVGGVTAYFRGVQAR 294

Db 94 KGLY---SGLAGNIAGVLPASALFVGVEPTKOKLLKTPPD--HLSAVAHLTAGAIGGLA 148
QY 82 TLLHDAAMNPAEVLKQRMOMYNPSYHRVTDVAVWQNECAGAFYRSYTTQLTMTNVPFQA 141
Db 149 ASL---IRVPTVVVKORMO--TGQFTSAPSASVVRMIASKEGFRGLYAGYRSFLLRDLFPDA 203
QY 142 IHFMVYEFLOEHFN--PQRRYNPSSHVLSGACAGAVAAAATPLDVCKTLLNTQESLALN 199
Db 204 IQFCIYEQLCLGYKKAARRELSDPENALIGAFAGALTGAVTTPLDVKTFLMWQGS---- 259
QY 200 SHITGHITGMASAFRTVYQVGGVTAYFR 227
Db 260 ---AKOYQIVDCVQTVIREEGAPALLK 284

Search completed: August 27, 2002, 03:57:42
Job time: 4948 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:16:46 ; Search time 37.5 Seconds
(without alignments)
273.618 Million cell updates/sec

Title: US-09-870-113-8
Perfect score: 1406
Sequence: 1 MQLQDPAPARYRNVLALW.....VYEFKYLITKROEWAGK 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 3871950 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	441	31.4	304	1 MRS4_YEAST	P23500 saccharomyc
2	428	30.4	314	1 MRS3_YEAST	P10566 saccharomyc
3	283	20.1	328	1 Y051_CAEEL	Q09461 caenorhabdi
4	275.5	19.6	368	1 YK39_YEAST	Q03829 saccharomyc
5	251.5	17.9	359	1 CG69_HUMAN	Q9bzj4 homo sapien
6	226	16.1	284	1 PEF8_YEAST	P38921 saccharomyc
7	217	15.4	676	1 CMC2_MOUSE	Q9qxx4 mus musculu
8	213	15.1	320	1 CMC2_HUMAN	Q9huc21 homo sapien
9	212.5	15.1	320	1 DMC_HUMAN	Q9ujs0 homo sapien
10	210	14.9	695	1 CMC1_DROME	Q9va73 drosophila
11	209.5	14.9	436	1 B71_MAIZE	P29518 zea mays (m
12	208	14.8	366	1 YG5F_YEAST	P53320 saccharomyc
13	206	14.7	322	1 SFCL_YEAST	P33303 saccharomyc
14	204	14.5	271	1 YAD8_SCHPO	Q09834 schizosacch
15	193.5	14.2	307	1 ODC2_YEAST	Q99297 saccharomyc
16	198	14.1	345	1 YDE9_SCHPO	Q10442 schizosacch
17	196.5	14.0	315	1 MFT_HUMAN	Q9h2d1 homo sapien
18	191	13.6	588	1 CMC2_CAEEL	Q20799 caenorhabdi
19	189.5	13.5	301	1 ORT1_MOUSE	Q9wvd5 mus musculu
20	187	13.3	678	1 CMC1_HUMAN	Q97546 homo sapien
21	186	13.2	315	1 SA18_HUMAN	Q9h1k4 homo sapien
22	186	13.2	373	1 YIA6_YEAST	P40556 saccharomyc
23	184	13.1	300	1 YEO3_YEAST	P40035 saccharomyc
24	184	13.1	335	1 YEA6_YEAST	P39953 saccharomyc
25	183.5	13.1	301	1 MCAT_HUMAN	Q43772 homo sapien
26	183	13.0	312	1 SA18_MOUSE	Q9db41 mus musculu
27	176	12.5	301	1 MCAT_RAT	P97521 rattus norv
28	174.5	12.4	301	1 ORT2_HUMAN	Q9bxi2 homo sapien
29	174.5	12.4	324	1 PMT_YEAST	P32332 saccharomyc
30	173	12.3	330	1 GDC_BOVIN	Q01888 bos taurus
31	172	12.2	587	1 CMC3_CAEEL	Q19529 caenorhabdi
32	170.5	12.1	311	1 UCP3_CANFA	Q9n219 canis famil
33	168.5	12.0	302	1 YD1K_SCHPO	Q10248 schizosacch

ALIGNMENTS

RESULT 1	MRS4_YEAST	STANDARD;	PRT;	304 AA.
ID	MRS4_YEAST			
AC	P23500;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Mitochondrial RNA splicing protein MSR4.			
GN	MSR4 OR YKR052C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=M1301.			
RC	MEDLINE=91108615; PubMed-1703236;			
RA	Wiesenberger G., Link T.A., von Ahlsen U., Waldherr M., Schweyen R.J.;			
RT	"MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast,			
RT	are new members of the mitochondrial carrier family.";			
RL	J. Mol. Biol. 217:23-37(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Visiers S., Urrestazu L.A., Jauniaux J.-C.;			
RL	Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: MRS4 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE			
CC	FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING			
CC	ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN			
CC	THE MITOCHONDRION (POSSIBLY OF CATIONS). NOT ESSENTIAL.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial			
CC	inner membrane.			
CC	-!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,			
CC	TO YEAST MRS3; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER			
CC	PROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X56444; CAA39828.1; -			
DR	EMBL; Z28277; CAA82130.1; -			
DR	PIR; S13533; S13533.			
DR	SGD; S0001760; MRS4.			
DR	InterPro; IPR001993; Mitoch_carrier.			
DR	Pfam; PF00153; mito_carr; 3.			
DR	PROSITE; PS00215; MITOCH_CARRIER; 2.			
KW	Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;			
KW	mRNA processing.			
FT	REPEAT 9 110 BLOCK I (APPROXIMATE).			
FT	REPEAT 111 202 BLOCK II (APPROXIMATE).			
FT	REPEAT 203 304 BLOCK III (APPROXIMATE).			
FT	TRANSMEM 23 41 POTENTIAL.			

P56501 mus musculu
P53007 homo sapien
P16260 homo sapien
P32089 rattus norv
P79110 bos taurus
P34519 caenorhabdi
Q9y619 homo sapien
O97649 sus scrofa
P56499 rattus norv
Q12375 saccharomyc
Q21153 caenorhabdi
P23641 saccharomyc

1 UCP3_MOUSE
1 TXTP_HUMAN
1 GDC_HUMAN
1 TXTP_RAT
1 TXTP_BOVIN
1 TXTP_CAEEL
1 ORT1_HUMAN
1 UCP3_PIG
1 UCP3_RAT
1 ORT1_YEAST
1 CMC1_CAEEL
1 MPCP_YEAST

34 168.5 12.0 308
35 168 11.9 311
36 167 11.9 332
37 166.5 11.8 311
38 165.5 11.8 311
39 165 11.7 312
40 164.5 11.7 301
41 164.5 11.7 308
42 164.5 11.7 308
43 163 11.6 292
44 163 11.6 702
45 162.5 11.6 311


```
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier C16C10.1.
GN C16C10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Lloyd C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; Z46787; CAA86739.1; -
DR WormPep; C16C10.1; CE01489.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Repeat; Transmembrane;
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 328 AA; 36743 MW; EAL9E329A764DF6 CRC64;

Query Match 20.1%; Score 283; DB 1; Length 328;
Best Local Similarity 30.9%; Pred. No. 4.9e-19;
Matches 82; Conservative 39; Mismatches 112; Indels 32; Gaps 7;

QY 12 YRVNLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKTKLSDVI----- 64
DB 74 FRGTADAIVKTAHEGIRSLSGSLPTVMALPATVFTTYDNLVSVLKKMKCCRRAPS 133

QY 65 ----HPGNSHIANAGACVATLLHDAAMNPAEVVQKQMNSPYHRVTCVRAVQNE 120
DB 134 PEKWTTPDWS--AAVAGIVARTIAVTVVSPLEMTITKMSKRLTYHEIGHVLRSMATK 191

QY 121 GAGAFYRSYTTQLTMMNVPQAIHEFTYEFLOEHFNPQRPNPSSH-----VLSGACAGAV 175
DB 192 GISSPYLGWTPTMLRDIPFGSIYAGYDLFKT--NLQRQGDHNPVSVFVSGAAAGVV 249

QY 176 AAATATPLDVCKTLNLTQESLALNSHITGTMASAFRTV-----YQGGVTAYFRGVA 231
DB 250 ASIFTHPFDVTKT--NCO-----IRIGSIDDMNKSITTIKDMYHSGISAFSSGLVP 301

QY 232 RVVIQIPSTAIWSVYEFKYLITK 256
DB 302 RLKRVSPSCAIMSEYEFKFLQK 326

RESULT 4
YM39_YEAST
ID YM39_YEAST STANDARD; PRT; 368 AA.
AC O03829;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative mitochondrial carrier YMR166C.
GN YMR166C OR YM8520.15C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
```

```
RN
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Hunt S.; Bowman S.; Barrell B.G.; Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; Z49705; CAA89802.1; -
DR SGD; S0004776; YMR166C.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
SQ SEQUENCE 368 AA; 40992 MW; B583100018DF045D CRC64;

Query Match 19.6%; Score 275.5; DB 1; Length 368;
Best Local Similarity 28.6%; Pred. No. 2.8e-18;
Matches 77; Conservative 59; Mismatches 100; Indels 33; Gaps 9;

QY 11 YRVNLEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKTKLSDVIHPGN 69
DB 89 KYRMISAVRTIWEEGVRRLGYGYMAALGSPFSAIFFGTVEYTKRTMEDQI--N 146

QY 70 SHIANGACCVATLLHDAAMNPAEVVQKQM-----YHRVTCVRAVQWQ 119
DB 147 DTITLSAGFLGDFTSFVYVPEVLKTRLQGLQGNPNPFQSGYNSLRLNAITVKE 206

QY 120 EGAGAFYRSYTTQLTMMNVPQAIHEFTYEFLOEHFNPQR-----YNPSSHVLSGACA 172
DB 207 EGRSLFPGYKATLARDLPFSALQFAFYEFKQLAFKLEQKDRGELSIPNEILTGACA 266

QY 173 GAVAAAAATPLDVCKTLNLTQESLALNS-----HIT-GHITGMAS-----AFRTVYQV 219
DB 267 GGLAGIITPMDVTRVQTPQPPSQSNKSYSVTHPHVTNGRPAALSNSISLRLTVYQS 326

QY 220 GGVTAIFRGVQARVYIQIPSTAIANSVYE 248
DB 327 EGVLGFFSGVGRFVMTSVQSSIMLLLYQ 355

RESULT 5
CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9BZJ4; Q9UF66; Q9Y379; Q9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
```

RA Yu X.X., Levin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
 RA Pan G., Adams S.H.;
 RT "Overexpression of the human 2-oxoglutarate carrier lowers
 RT mitochondrial membrane potential in HEK-293 cells: contrast with the
 RT unique cold-induced mitochondrial carrier CGI-69.";
 RL Biochem. J. 353:369-375(2001).
 RN [2]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=2072150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 RT Caenorhabditis elegans by comparative proteomics.";
 RL Genome Res. 10:703-713(2000).
 RN [3]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Colon;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN
 RP SEQUENCE OF 15-359 FROM N.A. (ISOFORM 1).
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/CGI-69L (shown here) and 2;
 CC are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in many tissues. Abundant in testis
 CC and kidney.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF317711; AAG60687.1; -
 DR EMBL; AF151827; RAD34064.1; -
 DR EMBL; AL133584; CAB63728.1; -
 DR EMBL; AK026060; BAB15341.1; -
 DR EMBL; BC001398; AAH01398.1; -
 DR EMBL; BC009330; AAH09330.1; -
 DR EMBL; AF119864; AAF69618.1; -
 DR InterPro; IPR001993; MitochCarrier.
 DR InterPro; IPR002067; MitCarrier.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.

DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Alternative splicing; Polymorphism.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT VARSPLIC 64 72 LPSSLOSTG -> W (IN ISOFORM 2).
 FT VARIANT 247 247 L -> F. /FTIG=VAR_012756.
 FT CONFLICT 266 266 G -> R (IN REF. 6).
 SQ SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

 Query Match 17.9%; Score 251.5; DB 1; Length 359;
 Best Local Similarity 27.8%; Pred. No. 4.7e-16;
 Matches 75; Conservative 43; Mismatches 121; Indels 31; Gaps 8;

 QY 7 DPAAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACVEKLKKTLL---60
 DB 100 DP-TRETGTMDAFVKIVRHGRTFLW---SGLPATLVMTPATAIYFTAYDQLKAFLCGR 155
 QY 61 ---SDVIHPGNGSHIANGAGCVATLLHDAAMNPAEVKQRMOMYNSPHRYVDCVRVW 117
 DB 156 ALTSLEYAP-----MVAGALARLGTVTVISPLELMRTKLQAQHVSYRELGCACVRTAV 207
 QY 118 ONEGAGAFYSYTTQLTMTNVPQAIHFMYEFLOEHFN---PQRYNPNSSHVLSGACAGA 174
 DB 208 AOGGWRSLWGLMGPTALRDVPFSALYWFYELVKSLWNLGRPKDQTSVGMFSFVAGGISGT 267
 QY 175 VAAATPTPLDVCKT---LLNTOESLALNSHITGHTGTMASAPRTVYQGVGTAYFERGVQ 230
 DB 268 VAAVTLPLFDVVKTQROVALGAEVAVNPL---HYDSTWLLLRRAESGTGLFAGFL 324
 QY 231 ARVIQIPSTAIWSYVEFFKYLITKROEE 260
 DB 325 PRIKAAPSCAIMISTYERFGKSFQRLNQD 354

 RESULT 6
 ID PET8_YEAST STANDARD; PRT; 284 AA.
 AC P38921;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Putative mitochondrial carrier protein PET8.
 GN PET8 OR YNL003C OR N2012.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GRF88;
 RX MEDLINE=94005822; PubMed=8402262;
 RA Lalo D., Stettler S., Mariotte S., Slonimski P.P., Thuriaux P.;
 RT "Two yeast chromosomes are related by a fossil duplication of their
 RT centromeric regions.";
 RL C. R. Acad. Sci., III, Sci. Vie 316:367-373(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GRF88;
 RX MEDLINE=95028151; PubMed=7941739;
 RA Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;
 RT "Organization of the centromeric region of chromosome XIV in
 RT Saccharomyces cerevisiae.";
 RL Yeast 10:523-533(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;


```
Matches 68; Conservative 41; Mismatches 110; Indels 36; Gaps 8;
QY 12 YRNVLALWRIITRGLRPMRGLNVTATGAGPAHALYFACYEKLKTLSDVI-----HP 66
Db 371 YKNSDFCKKVIJURYEGFGLYRGLLPQLLGVAPAKA-----IKLTVNDVFRDKFMHK 422
QY 67 GGN-----SHIANGAAGCVATLLHDAAMPVAVKQRMQNSPYHRVDCVR-----AVW 117
Db 423 DGSVPLAAEILAGGAGGSGQVIF---TNPLEIVKIRLQVAG---EITGPRVSALS 474
QY 118 QNEGAGAFYRSYTTQLTNNVPQAIHFWYEFLOEHF-NPQRRYNPSSHVLSGACAGAVA 176
Db 475 RDLGFFGIYKGAACFLRDIPEFSAIYPCYAHVKASAFANEDQVSPGSLLAGAIGMFA 534
QY 177 AAATPLDVKCKLLNTQESLALNSHITGTMASAFRTVYGVGTAVFVGVQARVIYQ 236
Db 535 ASLVTADVIKIRLQVAVRAGQTTI-----SGVIDCFKRLKEGPKALWKAGARVFRS 589
QY 237 IPSTAIAMSVYEFFK 251
Db 590 SPOFGVTLTYELLQ 604
RESULT 9
DNC_HUMAN
ID DNC_HUMAN STANDARD; PRT; 320 AA.
AC Q9HC21;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Mitochondrial deoxynucleotide carrier (Mitochondrial uncoupling protein 1).
GN SLC25A19 OR DNC OR MUP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11226231;
RA Doice V., Fiermonte G., Runswick M.J., Palmieri F., Walker J.E.;
RT "The human mitochondrial deoxynucleotide carrier and its role in the
RT toxicity of nucleoside antivirals."
RL Proc. Natl. Acad. Sci. U.S.A. 98:2284-2288(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Renard S., Mondesert G., Besnard F.;
RT "MUP 1, a mitochondrial uncoupling protein."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Responsible for the uptake of deoxynucleotides into the
CC matrix of the mitochondria. Transports all four deoxy NDPs, and,
CC less efficiently, the corresponding dNTPs. Does not transport
CC dNMPs, NMPs, deoxynucleosides, nucleosides, purines, or
CC pyrimidines. Supply deoxynucleotides to the mitochondrial matrix
CC for conversion to triphosphates and incorporation into
CC mitochondrial DNA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined except for
CC placenta. Highest levels in colon, kidney, lung, testis, spleen,
CC and brain.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- DISEASE: Likely to be medically important by providing the means
CC of uptake into mitochondria of nucleoside analogs, leading to the
CC mitochondrial impairment that underlies the toxic side effects of
CC such drugs in the treatment of viral illnesses, including AIDS,
CC and in cancer therapy.
```

```
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AJ251857; CAC27560.1; -
DR EMBL; AF301616; CAC37793.1; -
DR EMBL; AF182404; AAG16903.1; -
DR EMBL; BC001075; AAH01075.1; -
DR EMBL; BC005120; AAH05120.1; -
DR MIM; 606521; -
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
SQ SEQUENCE 320 AA; 35511 MW; 57CE0F01D538B1BE CRC64;
Query Match 15.1%; Score 212.5; DB 1; Length 320;
Best Local Similarity 23.4%; Pred. No. 1.8e-12;
Matches 62; Conservative 54; Mismatches 116; Indels 33; Gaps 9;
QY 7 DPAAIRYNNVLEALRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKTLSDV 63
Db 54 DPSAKYHGILQASRQIQLEEGPTAFWKGHVPAQILSIGYGAQVQLSF-----EMLT 106
QY 64 IHPGQ-----NSH-IANGAAGCVATLLHDAAMPVAVKQRMQNSPYHRVDCV 113
Db 107 VHRGVSVDAREFSVHFVCGGLAACMATL---TVHPVDVLRTRFAAQGEKPKYNTLRHAV 162
QY 114 RAWNOGAGAFYRSYTTQLTNNVPQAIHFWYEFLOEHFN-----PQRRYNPSSHVLS 168
Db 163 GTMYRSEGQVQVYKGLAPTLIAIFPVAGLQFSCYSLKHLKYWAIPAEGCKKNENLQNLIC 222
QY 169 GACAGAVAAAATPLDVCKTLNT---QESLALNSHITGTMASAFRTVYGVGTAY 225
Db 223 GSGAGVISKTLTYPLDLFKRLQVGVGFHARAAGVRRY-KGLMDCAKQVQLQKEGALGF 281
QY 226 FRGQVQARVIYQIPSTAIAMSVYEFF 250
Db 282 FKGLSPLSLKAAALSTGFMEFFSYEFF 306
RESULT 10
CMCL_DROME
ID CMCL_DROME STANDARD; PRT; 695 AA.
AC Q9VA73; Q9VA72; Q9VA74; Q9U5V8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding mitochondrial carrier Atalar1.
GN ARALARI OR CG2139.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Ovary;
RX MEDLINE=20115110; PubMed=10642534;
```

RA Del Arco A., Agudo M., Satrústegui J.;
 RT "Characterization of a second member of the subfamily of calcium-
 binding mitochondrial carriers expressed in human non-excitable
 tissues.";
 RL Biochem. J. 345:725-732(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Cocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinartz K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Sheng H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN; MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; MAY BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y18197; CAB62169.1; -
 CC EMBL: AE003774; AAF57048.1; -
 CC EMBL: AE003774; AAF57049.1; -
 CC EMBL: AE003774; AAF57050.2; -
 CC FlyBase: FBgn0028646; aralar1.
 CC InterPro: IPR002048; EF-hand.
 CC InterPro: IPR002067; Mit_carrier.
 CC InterPro: IPR001993; Mitoch_carrier.

DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00153; mito_carr; 3.
 DR PRINTS; PR00926; MITOCARRIER.
 DR SMART; SM00054; Eph; 3.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Calcium-binding; Alternative splicing.
 FT TRANSMEM 346 363 1 (POTENTIAL).
 FT TRANSMEM 406 425 2 (POTENTIAL).
 FT TRANSMEM 449 462 3 (POTENTIAL).
 FT TRANSMEM 498 517 4 (POTENTIAL).
 FT TRANSMEM 537 554 5 (POTENTIAL).
 FT TRANSMEM 594 613 6 (POTENTIAL).
 FT CA_BIND 84 95 EF-HAND 1.
 FT CA_BIND 118 129 EF-HAND 2.
 FT DOMAIN 157 168 ANCESTRAL CALCIUM SITE 3.
 FT CA_BIND 189 200 EF-HAND 4.
 FT DOMAIN 225 227 3 APPROXIMATE TANDEM REPEATS.
 FT REPEAT 225 227 1.
 FT REPEAT 435 524 2.
 FT REPEAT 525 627 3.
 FT VARSPLIC 1 52 MISSING (IN ISOFORM 3).
 FT VARSPLIC 3 22 MHIPFPFWIPLPVARCOE -> LTKSLPN (IN
 FT ISOFORM 2).
 FT CONFLICT 356 357 GA -> AP (IN REF. 1).
 FT CONFLICT 556 556 V -> A (IN REF. 1).
 FT CONFLICT 653 653 V -> L (IN REF. 1).
 SQ SEQUENCE 695 AA; 76753 MW; 8EC93D92031F5B77 CRC64;

 Query Match 14.9%; Score 210; DB 1; Length 695;
 Best Local Similarity 28.0%; Pred. No. 7.8e-12;
 Matches 72; Conservative 36; Mismatches 113; Indels 36; Gaps 8;

 QY 12 YRNVLEALWRIIRTEGLWPMRCLNVTATGAGPAHALYFACYEKLKLTSLDVTH----- 65
 DB 384 YRNSWDFCKVVRHGFGLGLPQLMGVAPEKA-----IKLVNDLVDRDKLTK 435
 QY 66 ----PGNSHIANGAAGCVATLLHDAAMPVAVVVKRMQ----MYSNPHYRVVDCVAVV 117
 DB 436 KGNIPWAEVLAGGAGAGSOVVF---TNPLEIVKIRLOVAGEIAGSGSKIRASVVREL- 490
 QY 118 QNEGAGAFYSYVTTQTMNVVFOAIHFMYEFLOEHNFQRRYN-PSSHVLSGACAGAVA 176
 DB 491 ---GLFLYKRGARCLLRDVPFSAIYFPTVYHTKAMADKDGYNHPLTLAAGAIAGVPA 547
 QY 177 AATTPDLVCKTLLNTQESIALNSHLTGHTGMASAFRTVYQVGVYATFRGVQARVIYQ 236
 DB 548 ASLVTADVIKTRL---QVVARSGQTT--YTGVDATKIMAEEGPRAFWKGTAAARVFRS 602
 QY 237 IPSTAIWSVYEFKYL 253
 DB 603 SPQFGVTLTYELLQRL 619

 RESULT 11
 BTL_MAIZE STANDARD; PRT; 436 AA.
 ID BTL_MAIZE
 AC P29518;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brittle-1 protein, chloroplast precursor.
 GN BTL.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Arafidaceae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93005685; PubMed=1668652;

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94203187; PubMed=7908717;
RA Fernandez M., Fernandez E., Rodicio R.;
RT "ACR1, a gene encoding a protein related to mitochondrial carriers,
RT is essential for acetyl-CoA synthetase activity in Saccharomycetes
RT cerevisiae.";
RL Mol. Gen. Genet. 242:727-735(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX Ramezani Rad M., Kirchthath L., Hollenberg C.P.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98053465; PubMed=9395087;
RA Palmieri L., Lasorsa F.M., De Palma A., Palmieri F., Runswick M.J.,
RA Walker J.E.;
RT "Identification of the yeast ACR1 gene product as a succinate-fumarate
RT transporter essential for growth on ethanol or acetate.";
RL FEBS Lett. 417:114-118(1997).
CC -!- FUNCTION: Transports cytoplasmic succinate, derived from
CC isocitrate by the action of isocitrate lyase in the cytosol, into
CC the mitochondrial matrix in exchange for fumarate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
CC -!- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
CC BY GLUCOSE.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z25485; CA80973.1; -;
DR EMBL; Z49595; CA89624.1; -;
DR PIR; S36407; S36407.
DR PIR; S43280; S43280.
DR SGD; S0003856; SFC1.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Transmembrane; Transport; Repeat.
FT CONFLICT 300 322 VREHLENLGIFKNDTPKPLK -> RKGAFAQKIWYSRR
FT MTHQSQSH (IN REF. 1).
SQ SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;

Query Match 14.78; Score 206; DB 1; Length 322;
Best Local Similarity 26.68; Pred. No. 7.2e-12;
Matches 69; Conservative 35; Mismatches 107; Indels 48; Gaps 10;

Qy 22 IIRTEGLWRPRLGNTATGAGPAHALYFACYEKLKLTLDVHPGNGSHIANGAA 77
Db 62 IYQEGFLALYKGLGAVVIGIIPKMAIRFSSYEFYRTLILVNKESGIYST-GNTEFVGVA 120
Qy 78 GCVATLLHDAAMNPAEVVVKRMQ-----MYNSPYHRVDCVRAVWQNEGAGA 124
Db 121 GITEAVL---VYNEVEVVKIRLQAHLTPSPNAGPKYNNIAHAAYTVK-----EEGVSA 173
Qy 125 FYRSYTTQLTMNVFPQAIHFMTY----EFLQEHFNPPRRYPSHVSLSGACAGAVAAAA 179
Db 174 LYRGVSLTAARQAATNOGANFTVYSKLKFLQNYH--QMDVLPSTWETSIGLISGAIGPFS 231
Qy 180 TPLDVCKTLTNTQESALNSHITGHTGMASAFRTVYQVG-----GVTAFFRGVQAR 232
Db *232 NAPLDTITKLRQDKSISLEKQ-----SGMKIITIGLAQLKBEGRFALYKGTIPR 282

Qy 233 VIYQIPSTAIANSVYEFK 251
Db 283 VNRVAPGQAVTFTVYEVYR 301

RESULT 14
YAD8_SCHPO STANDARD; PRT; 271 AA.
AC Q09834;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Putative mitochondrial carrier C4G8.08.
GN SPAC48.08.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (Potential).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z56276; CA91209.1; -;
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
SQ SEQUENCE 271 AA; 30002 MW; EF330C8CC8BDB6C4 CRC64;

Query Match 14.58; Score 204; DB 1; Length 271;
Best Local Similarity 27.08; Pred. No. 8.9e-12;
Matches 61; Conservative 41; Mismatches 110; Indels 14; Gaps 5;

Qy 33 RGLNVATGAGPAHALYFACYEKLKLTLDVHPGNGSHIANGAAGCVATLLHDAAMNPA 92
Db 48 RGLPISLTLTPATCLYLSTVVEAKRRPKPSV---GEGAILYSGMTAEVVSFVWTPPL 104
Qy 93 EVVKRQMOMYSPYHRVDCVRAVWQNEGAFYRSYTTQLTMNVFPQAIHFMTYEFLOE 152
Db 105 EVIKARTQI--SQKGSVINTISTLARSEGLKGFYRGYWMGVAIYLPPTVSWWVCYESKK 162
Qy 153 HFNFORRYNPSHVLSGACAG---AVAAAATPLDVCKTLTNTQESALNSHITG----H 205
Db 163 YLQKSNWDIS--VIAPICSAIGTVVATTISTPLDIVKTRYQVATSSAMRKAETGLQAEK 220
Qy 206 ITGMASAFRTVYQGVGTAYFRGVQARVIYQIIPSTAIANSVYEFK 251
Db 221 ELGILEIAKLFLSKHVGKGFTRGLFTRMCIYMPSGMISMSVFESEFK 266

RESULT 15

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:22 ; Search time 130.92 Seconds
(without alignments)
350.165 Million cell updates/sec

Title: US-09-870-113-8
Perfect score: 1406
Sequence: 1 MSLLQPDPAARYNRVLEALW.....VYEFKYLITKQBEWRAGK 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rviro:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1406	100.0	364	4	Q96A46	Q96A46 homo sapien
2	1007.5	71.7	338	11	Q920G8	Q920G8 mus musculus
3	937	66.6	177	4	Q969G8	Q969G8 homo sapien
4	910	64.7	176	4	Q9H2J3	Q9H2J3 homo sapien
5	848.5	60.3	347	4	Q9NY22	Q9NY22 homo sapien
6	725.5	51.6	187	4	Q9P0J2	Q9P0J2 homo sapien
7	722.5	51.4	187	11	Q9CQG7	Q9CQG7 mus musculus
8	677	48.2	312	5	Q23125	Q23125 caenorhabdi
9	657.5	46.8	379	5	Q9VAY3	Q9VAY3 drosophila
10	643	45.7	303	5	Q94638	Q94638 onchocerca
11	594	42.2	380	5	Q9NHY6	Q9NHY6 drosophila
12	573	40.8	301	5	Q94634	Q94634 onchocerca
13	467	33.2	303	3	O14281	O14281 schizosacch
14	431.5	30.7	781	10	Q9LMJ6	Q9LMJ6 arabidopsis
15	429.5	30.5	289	10	O82049	O82049 ribes nigru
16	414	29.4	331	10	O64731	O64731 arabidopsis

17	280	19.9	155	4	Q969S1	Q969s1 homo sapien
18	276	19.6	182	11	Q912Y0	Q912y0 mus musculus
19	273.5	19.5	320	11	Q921P8	Q921p8 mus musculus
20	266.5	19.0	321	4	Q9BSK2	Q9bsk2 homo sapien
21	266	18.9	413	5	Q9VWF9	Q9vwf9 drosophila
22	261.5	18.6	358	5	Q9VQ37	Q9vq37 drosophila
23	261.5	18.6	365	5	Q9STJ5	Q9stj5 drosophila
24	251.5	17.9	351	4	Q9Y379	Q9y379 homo sapien
25	251.5	17.9	351	4	Q9UF66	Q9uf66 homo sapien
26	251.5	17.9	359	4	Q9BZJ4	Q9bzj4 homo sapien
27	250	17.8	307	4	O14589	O14589 homo sapien
28	248.5	17.7	367	10	Q9C910	Q9c910 arabidopsis
29	243.5	17.3	337	4	Q9P182	Q9p182 homo sapien
30	242.5	17.2	325	10	Q94AG6	Q94ag6 arabidopsis
31	242.5	17.2	384	5	O27244	O27244 caenorhabdi
32	242	17.2	338	4	Q9UHR1	Q9uhr1 homo sapien
33	240.5	17.1	412	10	Q9FHX2	Q9fmx2 arabidopsis
34	238	16.9	322	10	Q9FU82	Q9fu82 oryza sativ
35	237	16.9	449	5	Q9VI05	Q9vi05 drosophila
36	233.5	16.6	359	11	Q9D8K8	Q9d8k8 mus musculus
37	231	16.4	706	3	Q9C2K1	Q9c2k1 neurospora
38	230.5	16.4	311	4	Q96CQ1	Q96cq1 homo sapien
39	230	16.4	310	4	Q9NVN5	Q9nvn5 homo sapien
40	224.5	16.0	902	3	Q12482	Q12482 saccharomyc
41	223.5	15.9	311	11	Q922G0	Q922g0 mus musculus
42	223	15.9	478	10	Q9FLS8	Q9fls8 arabidopsis
43	222	15.8	271	3	O60029	O60029 ashbya goss
44	219.5	15.6	269	5	Q18934	Q18934 caenorhabdi
45	218	15.5	676	11	Q9DCF5	Q9dcf5 mus musculus

ALIGNMENTS

RESULT 1

ID	Q96A46	PRELIMINARY;	PRT;	364 AA.
AC	Q96A46;			
DC	01-DEC-2001 (TREMREL. 19, Created)			
DT	01-DEC-2001 (TREMREL. 19, Last sequence update)			
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)			
DE	MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE CARRIER SPLICASE VARIANT).			
GN	HMR3/4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21195335; PubMed=11297739;			
RA	Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,			
RA	Larsson C., Suomalainen A.;			
RT	"Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.";			
RL	FEBS Lett. 494:79-84(2001).			
DR	EMBL; AJ303077; CAC27996.1; ..			
DR	EMBL; AF327402; AKA49519.1; ..			
SQ	SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64; *			

Query Match 100.0%; Score 1406; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.8e-129;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSLLQPDPAARYNRVLEALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTL	60
Db	100	MSLLQPDPAARYNRVLEALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTL	159
QY	61	SDVHPGGNSHIANGAGCVATLLHDAANPAEVVKORMQYNSPYHRVTDVCRAVWQNE	120
Db	160	SDVHPGGNSHIANGAGCVATLLHDAANPAEVVKORMQYNSPYHRVTDVCRAVWQNE	219


```

ID Q9NY22 PRELIMINARY; PRT; 347 AA.
AC Q9NY22;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HT015 PROTEIN.
GN HT015.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human hypothalamus.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223466; AAF64141.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.2.
SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 60.3%; Score 848.5; DB 4; Length 347;
Best Local Similarity 68.4%; Pred. No. 7.5e-75;
Matches 154; Conservative 34; Mismatches 36; Indels 1; Gaps 1;

QY 1 MQSLQPPAARYNRVLEALNRITTEGLWRPRLGNTATGAPAHALYFACYEKLKRTL 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 73 MQSLSPDPKQAYTSIYGALKIMTEGFWRLGYNVIMGAPAHAMFYACYENMKRTL 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 SDVTHPGNSHIANGACVATLLHDAAMPAEVVKORMQYNPSYHRVTDVRAVQNE 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 133 NDVPHGNSHLANGAGSATLLHDAAMPAEVVKORMQYNPSYHRVTDVRAVQNE 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GAGAFYSYTTQTMNVPFQAIHPTMYEFLQEHNPORRNPSSHVLSGACAGAVAAAT 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 193 GLGAFYSYTTQTMNVPFQAIHPTMYEFLQEHNPORRNPSSHVLSGACAGAVAAAT 252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 TPLDVCKTLNTQESLALN-SHITGHTITGMSAFRTVYQGVGTA 224
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 253 TPLDVCKTLNTQESLALN-SHITGHTITGMSAFRTVYQGVGTA 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
Q9POJ2 PRELIMINARY; PRT; 187 AA.
AC Q9POJ2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RA Huang Q., Yu Y., Huang C., Jiang C., Ren S., Zhou J., Gu W., Xu S.,
RA Wang Y., Fu G., Chen Z., Han Z.;
RT "A novel gene expressed in human adrenal gland.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF155660; AAF67479.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 2.
DR PRINTS; PR00926; MITOCH_CARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.1.
SQ SEQUENCE 187 AA; 21125 MW; 26F760526F7DE21B CRC64;

Query Match 51.4%; Score 722.5; DB 11; Length 187;
Best Local Similarity 71.2%; Pred. No. 7.4e-63;
Matches 131; Conservative 30; Mismatches 22; Indels 1; Gaps 1;

```

```

Query Match 51.6%; Score 725.5; DB 4; Length 187;
Best Local Similarity 71.9%; Pred. No. 3.7e-63;
Matches 133; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 80 VATLLHDAAMNPAEVVKORMQYNPSYHRVTDVRAVQNEGAGAFYSYTTQTMNVPF 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1 MATLLHDAAMNPAEVVKORLQYNPSYHRVTDVRAVQNEGAGAFYSYTTQTMNVPF 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 QAIHPTMYEFLQEHNPORRNPSSHVLSGACAGAVAAATTPLDVCKTLNTQESLALN 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 61 QSIHPTMYEFLQEHNPORRNPSSHVLSGACAGAVAAATTPLDVCKTLNTQESLALN 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 200 -SHITGHTITGMSAFRTVYQGVGTAAYFRGVAQVYQIPSTALMSVYEFKFLIKRQ 258
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 121 LANISGRLSGMANAFRTVYQGLAGYFKGIQARVITQMPSTALMSVYEFKFLIKRQ 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 EEWRA 263
   |||||
DQ 181 LENRA 185
   |||||

RESULT 7
Q9CQG7 PRELIMINARY; PRT; 187 AA.
AC Q9CQG7;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 1700020E22RIK PROTEIN.
GN 1700020E22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK015770; BAB29969.1; -.
DR EMBL; AK006155; BAB24436.1; -.
DR MGI; MGI:1914962; 1700020E22RIK.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 2.
DR PRINTS; PR00926; MITOCH_CARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.1.
SQ SEQUENCE 187 AA; 21194 MW; 8FB23583A8DD5E90 CRC64;

Query Match 51.4%; Score 722.5; DB 11; Length 187;
Best Local Similarity 71.2%; Pred. No. 7.4e-63;
Matches 131; Conservative 30; Mismatches 22; Indels 1; Gaps 1;

```

```

QY 80 VATLLHDAAMNPAEVYKQRMQYNSPYHRVTCVRAVWQNEGAGAFYRSTYTTQLTMNVFP 139
Db 1 MATLLHDAVWNPAAEVYKQRLQWYNSQSAFSCIRTVWTEGLGAFYRSTYTTQLTMNIF 60

QY 140 QATHMTYEFLOHFNPPORRYNPSSHVLSGACAGAAVAAATPLDVCKTLNLTQESLALN 199
Db 61 QSHFTYEFLOHFNPPORRYNPSSHVLSGACAGAAVAAATPLDVCKTLNLTQENMALN 120

QY 200 -SHITGHITGMSAFRTYVQGVGTAFYRQVAVIYQIPSTAIWSVVEFFKYLITKQ 258
Db 121 LANVSGRLSGMANAFYVQLNGLAGYFKGIQARVIYQMPSTAIWSVVEFFKYLITKQ 180

QY 259 EEMR 262
Db 181 LENR 184

RESULT 8
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
DR EMBL: Z66521; CAA91399.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 48.2%; Score 677; DB 5; Length 312;
Best Local Similarity 50.6%; Pred. NO. 4.2e-58;
Matches 133; Conservative 47; Mismatches 73; Indels 10; Gaps 3;

QY 1 MOSLOPDAARYNNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 45 MOSLCPCEPTKPTPHSLMSLVKREGWLPURGVNAAGSMPAHALYFTYERKMGYL 104

QY 61 SDVIHPGNSH-IANGAGCVATLLHDAAMNPAEVYKQRMQYNSPYHRVTCVRAVWON 119
Db 105 TG--NSAGHSNTLAYGASGVATLLHDAIMNPAEVYKQRMQMAFSPYSSLECARCVNR 162

QY 120 EGAGAFYRSTYTTQLTMNVFPQAIHMTVEFLOHFNPPORRYNPSSHVLSGACAGAAVAAA 179
Db 163 EGGAFAFYRSTYTTQLAMNVPFQAIHMTVEFLOHFNPPORRYNPSSHVLSGAGLAGLAAL 222

QY 180 TPLDVCKTLNLTQESLALN-----HITGHITGMSAFRTYVQGVGTAFYRQVAVIY 232
Db 223 TTPMDCVKTVLTQAAADPANRIFLQARYRYRGISDAVTVIYQRLSGFSGLQAR 282

QY 233 VYIQTIPSTAIWSVVEFFKYLIT 255
Db 283 VIFQVPATALWSVVEFFKFLMS 305

```

```

RESULT 9
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CG4963 PROTEIN (GH09840P).
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dushin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -!- INNER MEMBRANE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC EMBL: AE003763; AAF56764.1; -.
DR EMBL: AY060268; AAL25307.1; -.
DR FlyBase: FBgn0039561; CG4963.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.

```

```
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PRO0926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 46.8%; Score 657.5; DB 5; Length 379;
Best Local Similarity 51.2%; Pred. No. 4.5e-56;
Matches 130; Conservative 35; Mismatches 76; Indels 13; Gaps 3;

QY 1 MOSLQDPDPAARYRNVLALRIITRTGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MQSL--SPPTKNNIVSTLTREGLLRPIRGASAVILGAGPAHSLYFAAYEMTKELT 99
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVDCVRAVMQNE 120
Db 100 AKTTSVRLNLYVISA---VATLIHDAISSPTDVIKORMQYNSPYTSVSCVRDIYKR 156
QY 121 GAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNQRRYNPSSHVLSGACAGAVAAAAT 180
Db 157 GFRAFYRAYGTQLVNLPYQTIHFTYEFQKNKLERKYNPPVHMAAGAAAGACAAAAT 216
QY 181 TPLDVCKTLTNTQESLALNSHITGTMASAFRTYQVGGVTAAYFRGVQARVIYQIPST 240
Db 217 TPLDVTKTLTNTQE-----TGLTRGMIEASRKIIYHMAGPLGFFRGTTARVLYSMPT 268
QY 241 AIAWSVYEFKYL 254
Db 269 AICWSYIEFFKYL 282

RESULT 10
Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_Carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 45.7%; Score 643; DB 5; Length 303;
Best Local Similarity 46.4%; Pred. No. 8.7e-55;
Matches 122; Conservative 48; Mismatches 83; Indels 10; Gaps 2;

QY 1 MOSLQDPDPAARYRNVLALRIITRTGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 LQSLCPETSCPTAMHSLMSWKRGLLSLKGVNAVILGTIPARAFYTVTVENS KAYL 101
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKORMQYNSPYHRVDCVRAVMQNE 120
Db 102 LN--NPRVNSVSVAISGALATVHDVAMNPAEVVKORMQYNSPYGNSLECIRCIYRE 159
QY 121 GAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNQRRYNPSSHVLSGACAGAVAAAAT 180
Db 160 GLAFAYRSYTTQLTMMNVPFOAIHFMTYEFQKNKLERKYNPPVHMAAGAAAGACAAAAT 219

Query Match 42.2%; Score 594; DB 5; Length 380;
Best Local Similarity 47.5%; Pred. No. 7.5e-50;
Matches 121; Conservative 35; Mismatches 85; Indels 14; Gaps 4;

QY 1 MOSLQDPDPAARYRNVLALRIITRTGLWRPMRGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MQSL--SPPTKNNIVSTLTREGLLRPIRGASAVILGAGPAHSLYFAAYEMTKELT 99
QY 61 SDVIHPGGNSHIANGAGCVATLLHDAAMNPAEVVKOR--MQMYSNPYHRVDCVRAVMQ 119
Db 100 AKTTSVRLNLYVISA---VATLIHDAISSPTDVIKAYADCTTRPHTPVVSCVRDIYKR 156
QY 120 EGAGAFYRSYTTQLTMMNVPFOAIHFMTYEFLOEHFNQRRYNPSSHVLSGACAGAVAAA 179
Db 157 EGKAFAYRGTQLVNLPYQTIHFTYEFQKNKLERKYNPPVHMAAGAAAGACAAAAT 216
QY 180 TPLDVCKTLTNTQESLALNSHITGTMASAFRTYQVGGVTAAYFRGVQARVIYQIP 239
Db 217 TPLDVTKTLTNTQE-----TGLTRGMIEASRKIIYHMAGPLGFFRGTTARVLYSMPA 268
QY 240 TAIWSVYEFKYL 254
Db 269 TALCWSYIEFFKYL 283

RESULT 12
Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
```



```
QY 61 SDVIHPGG---NSHIANGAGCVATLLHDAAMPAEVVKQRMQMYNSPYHRVTDVCRAVWQ 118
| | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | | : : : :
Db 123 S-----AGDQNNVAHAMSGVFATISSDAVFTPMQVQKRLQMGEGTYKGWDCVKVLR 177
| | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | | : : : :
QY 119 NEGAGAFYSYTTQLTMMNVFQAIHMYEFLOE---HFNPRRYNPSS---HVLSGACA 172
| | | | | | | | : | | | | | | | | : | | | | | | | | : | | | | | | | |
Db 178 EEGIGAFYASYRTVLMNAPFTAVHFATYEAAKKGLMEFSPDRISDEGWLHVHATAGAA 237
| | | | | | | | : | | | | | | | | : | | | | | | | | : | | | | | | | |
QY 173 GAVAAAATTPDVKCTLLNTQESLALNSHITGHTGNASAFRTVYQVGGTAYFRGVQAR 232
| : | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 238 GGLAAAVTTPDVKVTKLQCCQCGCDRTSSSI---SHVLRITVKKDGYRGLRGWLPR 294
| : | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 233 VIYQIPSTAIAMSWYEFFKY 252
: : : | : | | | | | |
Db 295 MLFHAPAAAICWSTYEGVLY 314
: : : | : | | | | | |

RESULT 15
O82049
ID O82049 PRELIMINARY; PRT; 289 AA.
AC O82049;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MITOCHONDRIAL CARRIER PROTEIN.
GN PRIB7.
OS Ribes nigrum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Saxifragales; Grossulariaceae; Ribes.
OX NCBI_TaxID=78511;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
RA Woodhead M.R.;
RL Thesis (1995), University of Dundee, UK.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
RA Woodhead M.R.; Taylor M.A., Brennan R.M., McNicol R.J., Davies H.V.;
RT "Cloning and characterisation of the cDNA clones of five genes that
RT are differentially expressed during ripening in the fruit of
RT blackcurrant (Ribes nigrum L.).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -!- INNER MEMBRANE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC EMBL; AJ007580; CAA07568.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 289 AA; 30383 MW; 5666EAB03DCC507C CRC64;
```

```
Query Match 30.5%; Score 429.5; DB 10; Length 289;
Best Local Similarity 38.7%; Pred. No. 7e-34;
Matches 99; Conservative 38; Mismatches 108; Indels 11; Gaps 4;

QY 9 AARYRVNLEALWRIIRTEGLWRPMRGLNVNATGAGPAHALYFACYEKLKLTLSDVTHPGG 68
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 32 SAQSAGLRQALGSLKVEGPAGLYRGIGAMGLGAGPAHAVYFVSVMCKETFS---HGDP 88
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
QY 69 NSHIANGAGCVATLLHDAAMPAEVVKQRMQMYNSPYHRVTDVCRAVWQNEGAGAFYRS 128
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 89 SNSGAHAVSGVFATVADSVITPMQVQKRLQMGEGTYKGWDCVKVLRVVEEGIGAFYAS 148
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
QY 129 YTTQLTMMNVFQAIHMYEFLOE---HFNPRRYNPSSHVLSCACAVAAAATTP 183
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 149 YRTVVVWNPAPFTAVHFATYEAATKKLLEVPETANDENLLVHATAGAAAGALAAVVTPL 208
| : : | | : | | | | | | | | | | | | | | | | | | | | | | |
QY 184 DVCKTLLNTQESLALNSHITGHTGNASAFRTVYQVGGTAYFRGVQARVIYQIPSTAIA 243
```

```
Db 209 DVVKTLQCCQCGCDRTSSSIQDVIG---SIVKKNGYVGLMRGWTIPRMLFHAPAAAIC 265
| | | | | | | | : | | : | | : | | : | | : | | : | | : | | : | |
QY 244 WSYVEFFKYLITKRQE 259
| | | | | | | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 266 WSTYEASKTFFOKLINE 281
| | | | | | | | : | | : | | : | | : | | : | | : | | : | | : | |

Search completed: August 27, 2002, 04:19:23
Job time: 1290 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:55:59 ; Search time 138.54 Seconds
(without alignments)
75.364 Million cell updates

Title: US-09-870-113-10
 Perfect score: 502
 Sequence: 1 MOSLOPPAARYRNVLW.....GAAGCVATLLHDAAMNPAEG 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*

1:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #	Score	Query			ID	Description
			Match	Length	DB		
1	98.8	496	289	21	AA050393	Human uncoupling p	
2	98.8	496	289	22	AA060113	Human transport pr	
3	98.8	496	318	22	AA041505	Human polypeptide	
4	93.0	467	268	22	AA039719	Human polypeptide	
5	70.1	352	299	22	AA079039	Human protein SEQ	
6	69.9	351	272	21	AA042980	Human OREF ORF2744	
7	69.9	351	331	22	AA060658	Human mitochondria	
8	69.9	351	366	22	AA000938	Human bone marrow	
9	64.1	322	677	22	AA029748	Novel human secret	
10	55.0	276	155	21	AA042966	Human OREF ORF2730	
11	34.2	171.5	379	22	AA061130	Drosophila melanog	

12	165.5	33.0	59	22	AAM05891	Human immune/haema
13	130.5	26.0	289	18	AAW17034	Blackcurrant Rib7
14	130.5	25.9	278	21	AAG22079	Arabidopsis thalia
15	130	25.9	278	21	AAG43094	Arabidopsis thalia
16	130	25.9	290	21	AAG22078	Arabidopsis thalia
17	130	25.9	290	21	AAG43093	Arabidopsis thalia
18	130	25.9	331	21	AAG22077	Arabidopsis thalia
19	130	25.9	331	21	AAG43092	Arabidopsis thalia
20	121	24.1	130	22	AAM00825	Human bone marrow
21	99.5	19.8	413	22	ABB69108	Drosophila melanog
22	95.5	19.0	310	22	AAB93001	Human protein sequ
23	95.5	19.0	311	21	AAB50382	Human uncoupling p
24	95.5	19.0	311	22	AAM39173	Human polypeptide
25	95.5	19.0	374	22	AAM40959	Human polypeptide
26	95.5	19.0	374	22	AAM40960	Human polypeptide
27	94.5	18.8	237	21	AAG19645	Arabidopsis thalia
28	94.5	18.8	237	21	AAG50519	Arabidopsis thalia
29	94.5	18.8	241	21	AAG07172	Arabidopsis thalia
30	94.5	18.8	284	21	AAG19644	Arabidopsis thalia
31	94.5	18.8	284	21	AAG50518	Arabidopsis thalia
32	94.5	18.8	289	21	AAG07171	Arabidopsis thalia
33	94.5	18.8	296	21	AAG07170	Arabidopsis thalia
34	94.5	18.8	313	21	AAG19643	Arabidopsis thalia
35	94.5	18.8	313	21	AAG50517	Arabidopsis thalia
36	89.5	17.8	351	21	AAB50378	Human uncoupling p
37	89.5	17.8	365	21	AAB42783	Human ORF ORF2547
38	89	17.7	320	21	AAB50379	Human uncoupling p
39	89	17.7	320	22	AAM79071	Human protein SEQ
40	89	17.7	320	22	AAM93892	Human polypeptide,
41	89	17.7	329	22	AAM42411	Human polypeptide
42	89	17.7	329	22	AAM42413	Human polypeptide
43	89	17.7	335	21	AAB51795	Human secreted pro
44	87	17.5	358	22	ABB71142	Drosophila melanog
45	87	17.3	289	20	AAY31936	Wheat brittle-1 pa

ALIGNMENTS

RESULT	1
AAB50383	
ID	AAB50383 standard; Protein; 289 AA.
XX	
XX	
AAB50383;	
XX	
XX	
DT	12-MAR-2001 (first entry)
DE	Human uncoupling protein #6.

DR WPI; 2000-656322/63.
DR N-PSDB; AAC90457.
XX
PT Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
XX
PS Claim 11; Page 323-324; 343pp; English.
XX
CC The present sequence is a human uncoupling protein. The nucleotide
CC sequences encoding the uncoupling proteins may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
SQ Sequence 289 AA;

Query Match 98.8%; Score 496; DB 21; Length 289;
Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQSLQPPAARYRNVLALWRIIRTEGLWRPMSGINVTATGAGPAHALYFACYEKLKTL 60
Db 25 mqlqpdpaaaryrnvlealwriirteglwrpmsginvtatgagpahalyfacyeklkttl 84

Qy 61 SDVIHPGNSHIANGAGCVATLLHDAAMP 93
Db 85 sdvihpgnshiangagcvatllhdaampae 117

RESULT 2
AAB60113
ID AAB60113 standard; Protein; 289 AA.
XX
AC AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO200078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI; 2001-041424/05.
DR N-PSDB; AAF27733.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -

XX
PS Claim 2; Page 133-134; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTS). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 289 AA;

Query Match 98.8%; Score 496; DB 22; Length 289;
Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQSLQPPAARYRNVLALWRIIRTEGLWRPMSGINVTATGAGPAHALYFACYEKLKTL 60
Db 25 mqlqpdpaaaryrnvlealwriirteglwrpmsginvtatgagpahalyfacyeklkttl 84

Qy 61 SDVIHPGNSHIANGAGCVATLLHDAAMP 93
Db 85 sdvihpgnshiangagcvatllhdaampae 117

RESULT 3
AAM41505
ID AAM41505 standard; Protein; 318 AA.
XX
AC AAM41505;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6436.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60661.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6436; 10078pp; English.
XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AA42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 318 AA;

Query Match 98.8%; Score 496; DB 22; Length 318;
Best Local Similarity 100.0%; Pred. No. 4.9e-54;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWRPMLRGVLTATGAGPAHALYFACYEKLLKTL 60
DB 54 mslqdpdpaarynvlealwriirteglwrpmlrgvltatgagpahalyfacyekllkttl 113
QY 61 SDVIHPGNGSHANGAGCVATLLHDAAMNPAE 93
DB 114 sdvihpgngshangaagcvatllhdaampae 146

RESULT 4

AAM39719
ID AAM39719 standard; Protein; 268 AA.

XX AAM39719;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2864.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI: 2001-442253/47.
DR N-PSDB; AAI58875.

XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

XX Example 4; SEQ ID NO 2864; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AA42213) with neurotropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 268 AA;

Query Match 93.0%; Score 467; DB 22; Length 268;
Best Local Similarity 94.6%; Pred. No. 1.8e-50;
Matches 88; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWRPMLRGVLTATGAGPAHALYFACYEKLLKTL 60
DB 4 mslqsdpaarynvlealwriirteglwrpmlrgvltatgagpahalyfacyekllkttl 63

QY 61 SDVIHPGNGSHANGAGCVATLLHDAAMNPAE 93

DB 64 sdvihpgngshangaagcvatllhdaampae 96

RESULT 5

AAM79039

ID AAM79039 standard; Protein; 299 AA.

XX AAM79039;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1701.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB: AAK52172.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4043; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX Sequence 299 AA;

Query Match 70.1%; Score 352; DB 22; Length 299;
 Best Local Similarity 68.8%; Pred. No. 6.6e-36;
 Matches 64; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MOSLQDPAAARYNVLEALWRIITEGLWRPMRGVNTATGAGPAHALYFACYEKLKRTL 60
 Db 100 mqsispdpkadytsyvgalkkimrtgfwprlgrvnmimgagpahamyfacyenmkrtl 159
 Qy 61 SDVTHPGGNSHIANGAACGVATLLHDAAMNPAE 93
 Db 160 ndvfhgngshlangiagsgmatllhdavmnpae 192

RESULT 6

AAB42980
 ID AAB42980 standard; Protein; 272 AA.

XX AC AAB42980;

XX DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disease; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX W0200058473-A2.

XX PD

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAK77189.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 4662-4663; 5507pp; English.

XX AAK74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 272 AA;

Query Match 69.9%; Score 351; DB 21; Length 272;

Best Local Similarity 67.7%; Pred. No. 7.9e-36;

Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MOSLQDPAAARYNVLEALWRIITEGLWRPMRGVNTATGAGPAHALYFACYEKLKRTL 60

Db 7 mqsispdpkadytsyvgalkkimrtgfwprlgrvnmimgagpahamyfacyenmkrtl 66

Qy 61 SDVTHPGGNSHIANGAACGVATLLHDAAMNPAE 93

Db 67 ndvfhgngshlangiagsgmatllhdavmnpae 99

RESULT 7

AAB60658

ID AAB60658 standard; Protein; 331 AA.

XX AC AAB60658;

XX DT 04-MAY-2001 (first entry)

XX Human mitochondrial solute carrier protein hMSC-o.

XX Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
 KW preparation; detection.

```

XX OS Homo sapiens.
XX PN CN1269409-A.
XX PD 11-OCT-2000.
XX PF 17-MAR-2000; 2000CN-0114958.
XX PR 17-MAR-2000; 2000CN-0114958.
XX PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX PI Zhang X, Gao X, Xiao H;
XX DR WPI: 2001-050544/07.
XX DR N-PSDB; AAF59920.
XX PT New human mitochondrion solute carrier protein and its nucleic acid -
XX PS Claim 4; Page 20; 21pp; Chinese.
XX CC The invention relates to a novel human mitochondrial solute carrier
CC protein, hMSC-o (AAF60658), and cDNA encoding it (AAF59920). hMSC-o is
CC expressed in normal human hypothalamus tissue. The invention also relates
CC to the preparation of hMSC-o proteins and nucleic acids, and the
CC detection of hMSC-o proteins and nucleic acids in a sample. The present
CC sequence represents hMSC-o.
XX SQ Sequence 331 AA;

Query Match 69.9%; Score 351; DB 22; Length 331;
Best Local Similarity 67.7%; Pred. No. 1e-35;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPDPAARYNNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db ||||| ||| : : || : ||||| ||||| ||||| ||||| ||||| : ||
57 mqslspdkaqytsiygalkkimrtgfwprlgvnmimgagpahamyfacyenmkrtl 116
QY 61 SDVIHPGGNSHTANGAAGCVATILLHDAANPAE 93
Db : || : ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 ndvfhhqgnshlangiagsmatilhdavnmpae 149

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX AC AAM00938;
XX DT 01-OCT-2001 (first entry)
XX DE Human bone marrow protein, SEQ ID NO: 414.
XX KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX OS Homo sapiens.
XX PN WO200153453-A2.
XX PD 26-JUL-2001.
XX PF 23-DEC-2000; 2000WO-US34960.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.

```

```

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
XX DR N-PSDB; AAH90057.
XX PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX PS Claim 10; Page 504-505; 648pp; English.
XX CC The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX SQ Sequence 366 AA;

Query Match 69.9%; Score 351; DB 22; Length 366;
Best Local Similarity 67.7%; Pred. No. 1.1e-35;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPDPAARYNNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db ||||| ||| : : || : ||||| ||||| ||||| ||||| ||||| : ||
101 mqslspdkaqytsiygalkkimrtgfwprlgvnmimgagpahamyfacyenmkrtl 160

QY 61 SDVIHPGGNSHTANGAAGCVATILLHDAANPAE 93
Db : || : ||||| ||||| ||||| ||||| ||||| ||||| |||||
161 ndvfhhqgnshlangiagsmatilhdavnmpae 193

RESULT 9
AAU29748
ID AAU29748 standard; Protein; 677 AA.
XX AC AAU29748;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #239.
XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US08656.
XX PR 18-APR-2000; 2000US-0552929.

```

PR 26-JAN-2001; 2001US-0770160.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 DR
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy -
 PT
 XX Claim 20; Page 189; 765pp; English.
 PS
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 677 AA;

 Query Match 64.1%; Score 322; DB 21; Length 677;
 Best Local Similarity 63.4%; Pred. No. 1.1e-31; Indels 0; Gaps 0;
 Matches 59; Conservative 13; Mismatches 21;

 QY 1 MQSLQPDPAARYNRVLEALWRIIRTEGLWRPMLNLTATGAGPAHALYFACYEKLKTL 60
 Db 85 mqslspdqspvpsiygalkkimrtgfwrlrgvnmimgagpahamyfacyenmkrtl 144
 QY 61 SDVTHPGGNSHIANGAAGCVATLHDDAMNPAE 93
 Db 145 ndvfhggnshlangiagsmatllhdavmnpae 177

 RESULT 10
 AAB42966
 ID AAB42966 standard; Protein; 155 AA.
 XX
 AC AAB42966;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2730 polypeptide sequence SEQ ID NO:5460.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.

XX WO200058473-A2.
 XX
 XX 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach M;
 PI
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC77175.
 DR
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PT
 XX Claim 11; Page 4641; 5507pp; English.
 PS
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 155 AA;

 Query Match 55.0%; Score 276; DB 21; Length 155;
 Best Local Similarity 64.0%; Pred. No. 1.1e-26;
 Matches 48; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

 QY 1 MQSLQPDPAARYNRVLEALWRIIRTEGLWRPMLNLTATGAGPAHALYFACYEKLKTL 60
 Db 73 mqslspdqspqytsiygalkkimrtgfwrlrgvnmimgagpahamyfacyenmkrtl 132
 QY 61 SDVTHPGGNSHIANG 75
 Db 133 ndvfhggnshlang 147

 RESULT 11
 ABB61130
 ID ABB61130 standard; Protein; 379 AA.
 XX
 AC ABB61130;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 10182.
 XX

KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
XX	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPT; 2001-656860/75.
DR	N-PSDB; ABL05233.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Disclosure; SEQ ID NO 10182; 2lpp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 379 AA;
	Query Match 34.2%; Score 171.5; DB 22; Length 379;
	Best Local Similarity 40.9%; Pred.No.5.2e-13;
	Matches 38; Conservative 16; Mismatches 34; Indels 5; Gaps 2
QY	1 MQSLQPQAAARYRVLEALWRIIRTEGLWRPMRGINVTATGAGPAHALYFACYEKLKRTL 60
Db	: : : : :
	42 mqsl--spptkmmivstcrlmitrgllrpilrgasavvlgagpahslyfaayemtkelt 99
QY	61 SDVHPGNSHIANGAAGCVATLLHDAAMNPAE 93
Db	: :::: : :
	100 akfsvrnlrvyisga---vatlhdaissptd 129
RESULT 12	
AAM85891	ID
XX	AAM85891 standard; Protein; 59 AA.
AC	AAM85891;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:13484.
XX	
KW	Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	cytostatic; gene therapy; vaccine; metastasis.
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.


```
CC 57) that exhibit differential expression in blackcurrant fruit
CC during the ripening period of fruit development. RIB7 shows
CC 62% similarity to yeast MRS4, a yeast mitochondrial RNA splicing
CC protein. RIB7 is expressed almost entirely in fruit. The
CC promoter region (AAT68952) of the RIB7 gene can be used as a
CC fruit-specific promoter.
XX
SQ Sequence 289 AA;

Query Match 26.0%; Score 130.5; DB 18; Length 289;
Best Local Similarity 35.3%; Pred. No. 5.5e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

QY 9 AARYRNVLALMRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKTKTSLSDVTHPGG 68
   :|: : :||| :|: ||| :|: ||||| :|: ||| :|: |
Db 32 saqsaglrqalsilkevpaglrylganglgagpahavvfsvyemcketfs---hgdp 88

QY 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   :|: :|: ||| :|: ||| :|: |
Db 89 snsgahavsgvfatvasdavitpmd 113

RESULT 14
AAG22079
ID AAG22079 standard; Protein; 278 AA.
XX
AC AAG22079;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24871.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
```

```
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match      25.9%; Score 130; DB 21; Length 278;  
Best Local Similarity 37.9%; Pred. No. 6e-08;  
Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps  
  
QY   1 MQSLQDPDAARYRNVLKALWRIIRTEGLWRPMRGLNVTATGAGPAHALIFACYEKLKKTL 60  
    ||:||| : : : ||||:| | : |||||:| | | |  
Db   13 mqalrscp-ikpigirgafrsiiktgdpsalyrgiwamgigapahavfysfyevskkfl 71  
  
QY   61 SDVIHPGG--NSHIANGAAGCVATLLHDAAMPAE 93  
    | || |: |: |: ||| :| | :  
Db   72 s-----ggpnnsaaaisgvfatissdavftpmid 101  
  
RESULT 15  
AAG43094  
ID AAG43094 standard; Protein; 278 AA.  
XX AAG43094;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53822.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.
```

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.9%; Score 130; DB 21; Length 278;
Best Local Similarity 37.9%; Pred. No. 6e-08;
Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps 3;

Qy 1 MOSLQDPAAARYNNVLEALWRIIRTEGLWRPMPRGILNVTATGAGPAHALYFACYEKLKKTLL 60
Db ||:|: | : : | ||:|: | ||: |||||:|: || || |

Qy 13 mqlrscp-ikpigrqfrsiktgdpsalrygiwamglgagpahavvysfyevskkfl 71
Db ||:|: | : : | ||:|: | ||: |||||:|: || || |

Qy 61 SDVIHPGG--NSHIANGAAGCVATLLHDAAMNPAE 93
Db | || | : | : | ||: | | : | : |

Qy 72 s-----ggnpnsaahaaisgvfatissdgvftpm 101
Db | || | : | : | ||: | | : | : |

Search completed: August 27, 2002, 03:56:00
Job time: 4912 sec


```
Query Match          26.0%; Score 130.5; DB 4; Length 289;
Best Local Similarity 35.3%; Pred. No. 1.5e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

Qy 9 AARYRNVLALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSDVHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSLKVEGPAGLYRGICAMGLGAGPAHAVYFSVYEMCKETF5----HGDP 88

Qy 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNSGAHAVSGVFATVASDAVITPMD 113

RESULT 2
US-09-068-140A-15
; Sequence 15, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: NO. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-15

Query Match          26.0%; Score 130.5; DB 4; Length 328;
Best Local Similarity 35.3%; Pred. No. 1.8e-08;
Matches 30; Conservative 17; Mismatches 35; Indels 3; Gaps 1;

Qy 9 AARYRNVLALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKLTLSDVHPGG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 SAQSAGLRQALGSLKVEGPAGLYRGICAMGLGAGPAHAVYFSVYEMCKETF5----HGDP 88

Qy 69 NSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 SNSGAHAVSGVFATVASDAVITPMD 113

RESULT 3
US-08-933-750C-19
; Sequence 19, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNOAT01
; CLONE: 724157
US-08-933-750C-19

Query Match          17.8%; Score 89.5; DB 2; Length 351;
Best Local Similarity 32.3%; Pred. No. 0.0029;
Matches 31; Conservative 12; Mismatches 32; Indels 21; Gaps 5;

Qy 7 DPAARYRNVLALWRIIRTEG---LWRPRLGNVTATGAGPAHALYFACYEKLKLT--- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 DP-TRFTGTMDAFVKIVRHEGTRTLW---SGLPATLVMTVPATAIYETAYDQLKAFLCGR 147

Qy 61 ---SDVHPGNSHIANGAAGCVATLLHDAAMNPAE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 ALTSDLVAP-----MWAGALARLGTVTVISPLE 175
```



```

RESULT      4
US-09-234-613-19
; Sequence 19, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNOAT01
; CLONE: 724157
US-09-234-613-19

Query Match 17.8%; Score 89.5; DB 4;
Best Local Similarity 32.38%; Pred. No. 0.0029%; 32;
Matches 31; Conservative 12; Mismatches 32

QY 7 DPAARYNIVLEALWRIIRTEG---LWRPMRGLNVTATGAGPA
|| | : : | : | : | : | : | : | : | : | : | : |
DB 92 DP-TREFTGTMDFAKVIRHREGTRILW---SGLPATLVMTVPAN
|| | : : | : | : | : | : | : | : | : | : | : |
QY 61 ---SDVIHPGGNSHTANGAGCVATLLHDDAAMPAAE 93
|| | : | : | : | : | : | : | : | : | : | : |
DB 148 ALTSDLYAP-----MVAGALARLGTVTVISPLE 175

RESULT      5
US-08-933-750C-12
; Sequence 12, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:

```

```

> APPLICANT: Lal, Preeti
> APPLICANT: Hillman, Jennifer L.
> APPLICANT: Bandman, Olga
> APPLICANT: Shah, Purvi
> APPLICANT: Au-Young, Janice
> APPLICANT: Yue, Henry
> APPLICANT: Guegler, Karl J.
> APPLICANT: Corley, Neil C.
> TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
> NUMBER OF SEQUENCES: 98
> CORRESPONDENCE ADDRESS:
> ADDRESSEE: Incyte Pharmaceuticals, Inc.
> STREET: 3174 Porter Drive
> CITY: Palo Alto
> STATE: CA
> COUNTRY: USA
> ZIP: 94304
> COMPUTER READABLE FORM:
> MEDIUM TYPE: Diskette
> COMPUTER: IBM Compatible
> OPERATING SYSTEM: DOS
> SOFTWARE: FASTSEQ for Windows Version 2.0
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER: US/08/933.750C
> FILING DATE: September 23, 1997
> CLASSIFICATION: 536
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER:
> FILING DATE:
> ATTORNEY/AGENT INFORMATION:
> NAME: Billings, Lucy J.
> REGISTRATION NUMBER: 36,749
> REFERENCE/DOCKET NUMBER: PF-0356 US
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 415-855-0555
> TELEFAX: 415-845-4166
> TELEX:
> INFORMATION FOR SEQ ID NO: 12:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 320 amino acids
> TYPE: amino acid
> STRANDEDNESS: single
> TOPOLOGY: linear
> IMMEDIATE SOURCE:
> LIBRARY: SPLNOT02
> CLONE: 207452
> US-08-933-750C-12
>
Query Match 17.7%; Score 89; DB 2; I
Best Local Similarity 27.3%; Pred. No. 0.0029;
Matches 24; Conservative 18; Mismatches 28

Qy 7 DPAAARNVLEALWRIITEG---LWRPMRLNVTATGAGPAI
Db 54 DPSAKYHQLASRQILOEGPTAFWKGHVPQAQLILSIGYAV
Qy 64 IPFGG-----NSH-LANGAGCVATL 83
Db 107 VHRGSVDAREFSVHFVCGGLAACMATL 134

RESULT 6
US-09-234-613-12
> Sequence 12, Application US/09234613
> Patent No. 6132973
> GENERAL INFORMATION:
> APPLICANT: Lal, Preeti
> APPLICANT: Hillman, Jennifer L.
> APPLICANT: Bandman, Olga
> APPLICANT: Shah, Purvi
> APPLICANT: Au-Young, Janice
> APPLICANT: Yue, Henry

```

[illegible]

```

RESULT      6
US-09-234-613-12
; Sequence 12, Application US/09234613
; Patent No. 6132873
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-young, Janice
; APPLICANT: Yue, Henry

```

0
7
h
-
-
[
[
]

TELEPHONE: 609-243-0330

RESULT 10
US-08-937-466-4
; Sequence 4, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning

```

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1
```

```

Query Match          13.3%; Score 67; DB 3; Length 432;
Best Local Similarity 28.6%; Pred. No. 2.6; Indels 0; Gaps 0;
Matches 16; Conservative 11; Mismatches 29;

QY      8 PAARYRNVLALWRIIRTEGLWRPMRGLNVATGAGPAHALYFACYEKIKKTLSDV 63
        | ||| : : :: || : : : : : : : : || : || : || : |
Db     245 PLGYRSPHCLMKVAOEGPTAFYKGFPVFLRGAWNVMFMVTYEQLKRALMKV 300

RESULT 13
US-09-503-579-4
; Sequence 4, Application US/09503579
; Patent No. 6248561
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-503-579-4

Query Match          13.3%; Score 67; DB 4; Length 432;
Best Local Similarity 28.6%; Pred. No. 2.6; Indels 0; Gaps 0;
Matches 16; Conservative 11; Mismatches 29;

QY      8 PAARYRNVLALWRIIRTEGLWRPMRGLNVATGAGPAHALYFACYEKIKKTLSDV 63
        | ||| : : :: || : : : : : : : : || : || : || : |
Db     245 PLGYRSPHCLMKVAOEGPTAFYKGFPVFLRGAWNVMFMVTYEQLKRALMKV 300

RESULT 14
US-08-775-009-33
; Sequence 33, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weilong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
```

APPLICANT: Roe, Bruce
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
TITLE OF INVENTION: Velocardiofacial Syndrome Minimal Critical Region
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5935783ris, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-33

Query Match 13.0%; Score 65.5; DB 2; Length 311;
Best Local Similarity 30.6%; Pred. No. 2.6;
Matches 19; Conservative 12; Mismatches 28; Indels 3; Gaps 1;
QY 4 LQDPDPA---RYRNVLEALWRIIRTEGLWRPWRGLNVTATGAGPAHALYFACYEKLKLT 60
Db 53 LQIDERANPPRYRGIDCVQRQVIRSHVGLYRGLSSLLYGSIPRAAVRFGMFEFLSNEM 112
QY 61 SD 62
Db 113 RD 114

RESULT 15
US-09-268-347-32
Sequence 32, Application US/09268347
Patent No. 6335182
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268,347
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 1094
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-268-347-32

Query Match 12.9%; Score 65; DB 4; Length 1094;
Best Local Similarity 24.7%; Pred. No. 15;
Matches 22; Conservative 12; Mismatches 23; Indels 32; Gaps 3;

QY 34 GLNVTATGAGPAHALYFACYEKL---KKTLSVDVIHPGGNS-----H 71
Db 150 GVOVTSTSENGKHAITTFALAKDLDMTATVSDTLTIGGSTTTGSATPKVNVSTASGLN 209
QY 72 IANGAAGC-----VATLIHDAAMN 90
Db 210 FAKGATGANGDTTVVHLTNIASTLQDTLLN 238
Search completed: August 27, 2002, 03:58:53
Job time: 4895 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: August 27, 2002, 03:57:42 ; Search time 80.04 Seconds
(without alignments)
112.849 Million cell updates/sec
Title: US-09-870-113-10
Perfect score: 502
Sequence: 1 MOSLQDPDPAARYNVLEALW.....GAAGCVATLLHDAAMNPAEG 94
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_71:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	216.5	43.1	312	2 T26089	hypothetical prote
2	135	26.9	781	2 A86205	hypothetical prote
3	130	25.9	331	2 T00582	probable mitochond
4	124	24.7	314	2 S55179	mRNA splice defec
5	118	23.5	303	2 T39149	probable RNA splic
6	106	21.1	304	2 S13533	mRNA splice defec
7	101.5	20.2	384	1 S44092	probable carrier p
8	95	18.9	307	2 S60949	probable phosphate
9	94.5	18.8	296	2 B96830	hypothetical prote
10	94.5	18.8	313	2 D84613	hypothetical prote
11	93	18.5	436	2 JQ1459	Bcl protein precu
12	90	17.9	363	2 T49281	mitochondrial phos
13	89.5	17.8	351	2 T43493	hypothetical prote
14	89	17.7	348	2 D84798	probable mitochond
15	87.5	17.4	310	2 T20229	hypothetical prote
16	86.5	17.2	322	2 S57116	probable carrier p
17	85	16.9	311	2 G86383	probable mitochond
18	83.5	16.6	353	2 T51393	probable mitochond
19	83	16.5	284	2 T18253	probable mitochond
20	83	16.5	288	2 T51595	phosphate transpor
21	82	16.3	392	2 T05350	adenylate transloc
22	80.5	16.0	314	2 A36305	2-oxoglutarate/mal
23	79.5	15.8	296	2 T23170	hypothetical prote
24	79.5	15.8	314	2 A56650	2-oxoglutarate car
25	79	15.7	332	2 T47703	Ca-dependent solut
26	78.5	15.6	313	2 T05577	uncoupling protein
27	78.5	15.6	358	2 T45934	hypothetical prote
28	78	15.5	336	2 T01839	hypothetical prote
29	78	15.5	366	2 S64589	hypothetical prote

30	77.5	15.4	301	1 S31935	ADP,ATP carrier pr
31	77	15.3	325	2 E69118	lactate dehydrogen
32	77	15.3	368	2 S54524	hypothetical prote
33	76.5	15.2	298	2 A12615	formamidopyrimidin
34	76.5	15.2	325	2 H97397	formamidopyrimidin
35	76.5	15.2	1320	1 S66279	proline dehydrogen
36	76.5	15.2	1320	2 A50633	proline dehydrogen
37	76	15.1	282	2 T49628	probable dicarboxy
38	76	15.1	375	2 T05707	phosphate transpor
39	74.5	14.8	367	2 G96770	hypothetical prote
40	74.5	14.8	650	2 T32897	hypothetical prote
41	74	14.7	307	2 S67662	oxoglutarate/malat
42	73	14.5	290	2 S44091	hypothetical prote
43	73	14.5	418	2 B96811	phosphate transpor
44	72	14.3	366	2 T01169	phosphate transpor
45	71.5	14.2	330	2 T29640	mitochondrial carr

ALIGNMENTS

RESULT 1

T26089

hypothetical protein W02B12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26089

R:Swinsburne, J.; Ainscough, R.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z20149

A:Accession: T26089

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-312 <WILL>

A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9

A:Experimental source: clone W02B12

C:Genetics:

A:Map position: 2

A:Gene: CESP:W02B12.9

A:Introns: 18/3; 251/3; 286/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 43.1%; Score 216.5; DB 2; Length 312;

Best Local Similarity 50.0%; Pred. No. 2.2e-16;

Matches 47; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 MOSLQDPDPAARYNVLEALWRIIRTEGLWPMRGVLNVATGAGPAHALYFACYEKLKKTLL 60

DB 45 MOSLCPCPETKCTPTVHSLMSIVKREGWLRPLRGVNAAGSMPAHALYFTVYERMKGYL 104

QY 61 SDVIHPGNSH-IANGAACCVATLLHDAAMNPAE 93

DB 105 TG--NSAGHSNTLAYGASGVATLLHDAIMNPAE 136

RESULT 2

A86205

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A86205

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A86205
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-781 <STO>
 A:Cross-references: GB:AE005172; NID:g8954043; PIDN:AAF92217.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 26.9%; Score 135; DB 2; Length 781;
 Best Local Similarity 38.9%; Pred. No. 6.8e-07;
 Matches 37; Conservative 14; Mismatches 36; Indels 8; Gaps 3;

QY 1 MOSLOPDPARYRNVLALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKKTLL 60

Db 64 MQALRCP-LKPVGIREARFSIIQEGSPALYRGIVAMGLGAGPAHAYFSFEVSKKYL 122

QY 61 SDVIHPGG--NSHIANGAGCVATLLHDAAMP 93

Db 123 S-----AGDQNSVAHMSGVFATISSDAVFTPM 152

RESULT 3

T00582 probable mitochondrial carrier protein [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T27E13.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00582; C84705

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, May 1998

A:Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.

A:Reference number: Z14178

A:Accession: T00582

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-331 <ROU>

A:Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150404

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAren, S.E.; Umayam, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84705

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: GB:AE002093; NID:g3150404; PIDN:AAC16956.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g30160; T27E13.10

A:Map position: 2

A:Introns: 263/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

C:Keywords: mitochondrion

Query Match 25.9%; Score 130; DB 2; Length 331;

Best Local Similarity 37.9%; Pred. No. 9.7e-07;

Matches 36; Conservative 15; Mismatches 36; Indels 8; Gaps 3;

QY 1 MOSLOPDPARYRNVLALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKKTLL 60

Db 66 MQALRSCP-IKPTGIRQAFRSIIKTGPSALYRGIVAMGLGAGPAHAYFSFEVSKKFL 124

QY 61 SDVIHPGG--NSHIANGAGCVATLLHDAAMP 93

Db 125 S-----GGPNNSAAHISGVFATISSDAVFTPM 154

RESULT 4

S55179

mtRNA splice defect-suppressing mitochondrial carrier MRS3 - yeast (Saccharomyces cer
 N:Alternate names: protein J0675; protein YJL133W

C:Species: Saccharomyces cerevisiae

C>Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 21-Jul-2000

A:Accession: S55179; S01267; S20228; S56915; S71664

R:Katsoulou, C.; Tzermia, M.; Alexandraki, D.

submitted to the EMBL Data Library, May 1995

A:Description: The complete sequence of a 40.7 kb segment located on the left arm of

yeast hypothetical proteins.

A:Reference number: S55159

A:Accession: S55179

A:Molecule type: DNA

A:Residues: 1-314 <KAT>

A:Cross-references: EMBL:X87371; NID:g854542; PID:g854563

R:Schmidt, C.; Soellner, T.; Schweyen, R.J.

Mol. Gen. Genet. 210, 145-152, 1987

A:Title: Nuclear suppression of a mitochondrial RNA splice defect: nucleotide sequenc

A:Reference number: S01267; MUID:88121698

A:Accession: S01267

A:Molecule type: DNA

A:Residues: 1-86,'S',88-102,'L',104-127,'Y',129-141,'M',143-187,'R',189-314 <SCH>

A:Cross-references: EMBL:X06239; NID:g3990; PIDN:CAA29582.1; PID:g3991

R:Wiesenberger, G.; Link, T.A.; von Ahsen, U.; Waldherr, M.; Schweyen, R.J.

J. Mol. Biol. 217, 23-37, 1991

A:Title: MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast, are new m

A:Reference number: S13532; MUID:91108815

A:Accession: S20228

A:Molecule type: DNA

A:Residues: 19-314 <WIE>

A:Cross-references: EMBL:X56445; NID:g3992; PIDN:CAA39830.1; PID:g3994

R:Katsoulou, C.; Tzermia, M.; Alexandraki, D.

submitted to the Protein Sequence Database, September 1995

A:Reference number: S56912

A:Accession: S56915

A:Molecule type: DNA

A:Residues: 1-314 <KAW>

A:Cross-references: EMBL:Z49408; NID:gi1008337; PID:gi1008338; MIPS:YJL133W

R:Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.

Yeast 12, 787-797, 1996

A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome

of chromosome XI.

A:Reference number: S716643; MUID:96408771

A:Accession: S71664

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-314 <KAP>

A:Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60822.1; PID:g854563

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995

C:Genetics:

A:Gene: SGD:MRS3

A:Cross-references: SGD:S0003669; MIPS:YJL133W

A:Map position: 10L

A:Genome: nuclear

C:Function:

A:Description: probably involved in splicing of all intron for COB gene; essential fo

licing

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

C:Keywords: mitochondrion; inner membrane; mitochondrion; pre-mRNA splicing

F:30-119/Domain: ADP,ATP carrier protein repeat homology <ACP1>

F:127-211/Domain: ADP,ATP carrier protein repeat homology <ACP2>

F:216-311/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 24.7%; Score 124; DB 2; Length 314;

Best Local Similarity 36.7%; Pred. No. 4.3e-06;

Matches 29; Conservative 9; Mismatches 41; Indels 0; Gaps 0;

QY 13 RNVLALWRIIRTEGLWRPVRGLNVTATGAGPAHALYFACYEKLKKTLSVDIHPGNSHI 72

Db 72 KNMLSQISHISTSEGLALWKGVQSVILGAGPAHAYFGTYEFCCKNLIDSSDTQTHPF 131

QY 73 ANGAACVCATLLHDAAMP 91

Db 132 KTAISGACATTASDALMNP 150

RESULT 5
T39149
probable RNA splicing proteinmitochondrial carrier protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: T39149
R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21748
A;Accession: T39149
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-303 <OLI>
A;Cross-references: EMBL:Z99168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A;Experimental source: strain 972h-; cosmid c8C9
C;Genetics:
A;Gene: SPDB:SPAC8C9.12c
A;Map position: 1
C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

RESULT 6
 mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
 S13533
 N:Alternate names: protein YKR052c
 C:Species: Saccharomyces cerevisiae
 C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
 C:Accession: S13533; S38126
 R:Wiesenberger, G.; Link, T.A.; von Ahlsen, U.; Waidherr, M.; Schweyen, R.J.
 J. Mol. Biol. 217, 23-37, 1991
 A:Reference number: S13532; MUID:91108815
 A:Accession: S13533
 A:Molecule type: DNA
 A:Residues: 1-304 <MO>
 A:Cross-references: EMBL:X56444; NID:g3995; PIDN:CAA39828.1; PID:g3996
 R:Visiers, S.; Urrestarazu, L.A.; Jauniaux, J.C.
 Submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38118
 A:Accession: S38126
 A:Molecule type: DNA
 A:Residues: 1-304 <VIS>
 A:Cross-references: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:MRS4
 A:Cross-references: SGD:S0001760; MIPS:YKR052c
 A:Map position: 11R
 A:Genome: nuclear
 C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
 C:Keywords: duplication; mitochondrion; transmembrane protein
 F:20-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
 F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
 F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>
 Query Match 21.1%; Score 106; DB 2; Length 304;

```

Best Local Similarity 32.5%; Pred. No. 0.00041;
Matches 25; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

QY 15 VLEALWRIIRTEGLWRMRGLNVTATGAGPAHALYFACYEKLKLTSLDSVHPGGNSHTAN 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 MISQISKISTWEGSWALWKGVQSVILGAGPAHAYVFGTYEFCARLISPEDMQTHQPKMT 123

QY 75 GAAGCVATLLHDAAMNP 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 ALSGTIATIAADALMNP 140

RESULT 7
S44092
C:Probable carrier protein c2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C:Accession: S44092; T24753
R:Runswick, M.J.; Philipides, A.; Lauria, G.; Walker, J.E.

```

```

RESULT      7
S44092
C: probable carrier protein c2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C:Accession: S44092; T24753
R:Runswick, M.J.; Philipides, A.; Lauria, G.; Walker, J.E.
submitted to the EMBL Data Library, November 1993
A:Description: Extension of the mitochondrial transport superfamily: sequences
A:Reference number: S44090
A:Accession: S44092
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <RUN>
A:Cross-references: EMBL:X76116; NID:g472899; PIDN:CAA53722.1; PID:g472900
R:Lloyd, C.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19932
A:Accession: T24753
A>Status: preliminary; translated from GE/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-384 <WTL>
A:Cross-references: EMBL:Z49070; PIDN:CAA88869.1; GSPDB:GN00020; CESP:T09F3.2
A:Experimental source: Clone T09F3
C:Genetics:
A:Gene: CESP:T09F3.2
A:Map position: 2
A:Introns: 82/1; 113/1; 153/2; 179/2; 259/3; 359/3
C:Superfamily: Caenorhabditis probable carrier protein c2; ADP,ATP carrier prot
F:4-43,141-193/Domain: ADP,ATP carrier protein repeat homology #status atypical
F:45-67/Region: serine-rich
F:201-285/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:296-381/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match          20.2%; Score 101.5; DB 1; Length 384;
Best Local Similarity 35.1%; Pred. No. 0.0017;
Matches 33; Conservative 17; Mismatches 31; Indels 13; Gaps 5;

QY    5 QPDPAARYRN-VLEPALWKRIIRTEGLWRPMRLNVNTAGGAPAHALYFACYEKLKK--TLS 61
Db    137 QPTPAARGTIVIKITQVITIGEGALYGKLIPLNVGVAPSKAVIFYTTSTSRFWNES 196
               ||| ||| :::: ||:||||: || | | | | | | | | | | | | | | | |
QY    62 DVHPGNSHIAN----GAAGCVATLLHDAAAMP 91
               :: | | | | | :||| || | | | | | | | | | | | | | | | | |
Db    197 EVLPV--NSAIVHMVSAGSAGFVA-----SAVNP 224

RESULT      8
S60949
C: probable phosphate transport protein, mitochondrial - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5064; protein YOR222w; protein YOR50-12
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
C:Accession: S60949; S67115; B37984; S71724
R:Galisson, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of
A:Reference number: S60938
A:Accession: S60949
A:Molecule type: DNA

```

RESULT 8
S60949
probable phosphate transport protein, mitochondrial - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5064; protein YOR222w; protein YOR50-12
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
C:Accession: S60949; S67115; B37984; S71724
R:Galissou, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome III of *Saccharomyces cerevisiae*.
A:Reference number: S60938
A:Accession: S60949
A:Molecule type: DNA

Query Match 18.8%; Score 94.5; DB 2; Length 296;
Best Local Similarity 37.9%; Pred. No. 0.0077;
Matches 25; Conservative 8; Mismatches 20; Indels 13; Gaps 3;

QY 22 IIRTEGLWRPMRGLNVTATGAGPAHAHALFYFACYEKLLKTTISDVHFG-----GNSH-----I 72
| : | | || : | | | || | | | | : | | :
Db 150 ILRRQGQLGYRLGITIVLRDAPAHGLYFWTYEVRERL---HPGCRTKGQNLTMLV 205

QY 73 ANGAAG 78
| | | |
Db 206 AGGLAG 211

RESULT 10
D84613
hypothetical protein At2g22500 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84613
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84613
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 <STO>
A:Cross-references: GB:AE002093; NID:g544443; PIDN:AAD22351.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g22500
A:Map position: 2
C:Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology

Query Match 18.8%; Score 94.5; DB 2; Length 313;
Best Local Similarity 31.8%; Pred. No. 0.0081;
Matches 27; Conservative 23; Mismatches 28; Indels 7; Gaps 3;

QY 12 YRVLEALWRIITETGLWRPMRGLNTATGAGPAHAHALFYFACYEKLLKTLS-D-VIHPGCN 69
| ::|| |: ||| || : | | | | : | | : | | : | | : |
Db 161 YKSVDIAITOMIREGVTSILWRGSSTINRAMLVTSQASYSVKETILEXGLKLKDGLG 220

QY 70 SHI-ANGAACCVATLLHDAAANPAE 93
| : | : | | | | | | | : | | : |
Db 221 THVSASFAGFVASV----ASNPD 241

RESULT 11
JQ1459
Btl protein precursor - maize
C:Species: Zea mays (maize)
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Sep-1999
C:Accession: JQ1459
R/Sullivan, T.; Strelow, L.J.; Illingworth, C.A.; Phillips, R.L.; Nelson Jr., O.E.
Plant Cell 3, 1337-1348, 1991
A>Title: Analysis of maize brittle-1 alleles and a defective suppressor-mutator-induc
A:Reference number: JQ1459; MUID:93005685
A:Accession: JQ1459
A:Molecule type: mRNA
A:Residues: 1-436 <SUL>
A:Cross-references: GB:M79333; NID:q168425; PIDN:AAA33438.1; PID:q168426
C:Comment: This protein acts as an adenylate translocator in amyloplasts.
C:Genetics:
A:Introns: 206/3; 260/3
C:Superfamily: Btl protein; ADP, ATP carrier protein repeat homology
C:Keywords: chloroplast; duplication; transmembrane protein
F:1-75/DNA:transit peptide (amyloplast) #status predicted <TNP>
F:76-436/Product: Btl protein #status predicted <MAT>
F:131-217/Domain: ADP, ATP carrier protein repeat homology <ACPl>

Matches 27; Conservative 14; Mismatches 36; Indels 15; Gaps 3;

Qy 4 LOPDPA-ARYRNVEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKK---- 58

Db 125 IQDDKAHTKENGPIDATKQLLRTHGLKSLTRGFLATVARDAPAFGVYFASYEWARSCK 184

Qy 59 -----TLSDVIHPGGNSHIANGAAGCVATLLH 85

Db 185 DGETSTLS-----SGQLFAGGTAGMLSWLEN 211

Search completed: August 27, 2002, 03:57:43
Job time: 4949 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:16:48 ; Search time 37.5 Seconds
(without alignments)
97.057 Million cell updates/sec

Title: US-09-870-113-10
Perfect score: 502
Sequence: 1 MQSLQDPAPARYNVLEALW.....GAAGCVATLIHDAAMNPAEG 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	124	24.7	314	1 MRS3_YEAST	P10566 saccharomyc
2	106	21.1	304	1 MRS4_YEAST	P23500 saccharomyc
3	95	18.9	307	1 ODC2_YEAST	Q99297 saccharomyc
4	93	18.5	436	1 BFL1_MAIZE	P29518 zea mays (m
5	89.5	17.8	359	1 CG69_HUMAN	Q9bzi4 homo sapien
6	89	17.7	320	1 DNG_HUMAN	Q9hcz1 homo sapien
7	86.5	17.2	322	1 SFCL_YEAST	P33303 saccharomyc
8	80.5	16.0	313	1 M2OM_BOVIN	P22292 bos taurus
9	79.5	15.8	313	1 M2OM_HUMAN	Q02978 homo sapien
10	78	15.5	366	1 YG5F_YEAST	P53320 saccharomyc
11	77.5	15.4	301	1 ADTL_ANOGA	Q27238 anopheles g
12	77.5	15.4	313	1 M2OM_MOUSE	Q9cr62 mus musculu
13	77	15.3	368	1 YN39_YEAST	Q03829 saccharomyc
14	76.5	15.2	1320	1 PUTA_SALTY	P10503 salmonella
15	74.5	14.6	315	1 SA18_HUMAN	Q9h1k4 homo sapien
16	73.5	14.6	287	1 DIC_HUMAN	Q9ubx3 homo sapien
17	73.5	14.6	313	1 M2OM_RAT	P97700 rattus norv
18	71.5	14.2	315	1 MFT_HUMAN	Q9h2d1 homo sapien
19	71	14.1	308	1 UCP3_RAT	P36499 rattus norv
20	71	14.1	695	1 CMCL1_DROME	Q3va73 drosophila
21	70.5	14.0	287	1 DIC_MOUSE	Q9qzd8 mus musculu
22	70	13.9	308	1 ADT1_CHLRE	P27080 chlamydomon
23	69.5	13.8	1312	1 PUTA_KLEAE	O52485 klebsiella
24	69	13.7	301	1 PFG_RHINE	Q59752 rhizobium m
25	69	13.7	307	1 UCP1_HUMAN	P25874 homo sapien
26	69	13.7	345	1 YDE9_SCHPO	Q10442 schizosacch
27	68.5	13.6	312	1 UCP3_HUMAN	P55916 homo sapien
28	68.5	13.6	328	1 YQ51_CAEEL	Q09461 caenorhabdi
29	68	13.5	308	1 UCP3_PIG	O97649 sus scrofa
30	68	13.5	309	1 ADTL_YEAST	P40710 saccharomyc
31	68	13.5	381	1 ADTL1_ARATH	P31167 arabidopsis
32	67.5	13.4	311	1 UCP3_CANFA	Q9n219 canis famil
33	67	13.3	299	1 ODC_HUMAN	Q9bqt8 homo sapien

RESULT 1

ID	MRS3_YEAST	STANDARD;	PRT;	314 AA.
AC	P10566;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Mitochondrial RNA splicing protein MRS3.			
GN	MRS3 OR YJL133W OR J0675.			
OS	Saccharomycetes cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ICB/RJ01;			
RX	MEDLINE=88121698; PubMed=2448588;			
RA	Schmidt C., Soellner T., Schweyen R.J.;			
RT	"Nuclear suppression of a mitochondrial RNA splice defect: nucleotide sequence and disruption of the MRS3 gene.";			
RL	Mol. Gen. Genet. 210:145-152(1987).			
RN	[2]			
RP	REVISIONS, SEQUENCE FROM N.A.			
RC	STRAIN=M1301;			
RX	MEDLINE=91108815; PubMed=1703236;			
RA	Wiesenberg G., Link T.A., von Ahlsen U., Waldherr M., Schweyen R.J.;			
RT	"MRS3 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the mitochondrial carrier family.";			
RL	J. Mol. Biol. 217:23-37(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / FY1679;			
RX	MEDLINE=96408711; PubMed=8813765;			
RA	Katsoulou C., Termlia M., Tavernarakis N., Alexandraki D.;			
RT	"Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X reveals 14 known genes and 13 new open reading frames including homologs of genes clustered on the right arm of chromosome XI.";			
RL	Yeast 12:787-797(1996).			
CC	-1- FUNCTION: MRS3 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE			
CC	ITS SUPPRESSOR OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING IN			
CC	THE MITOCHONDRIUM (POSSIBLY OF CATIONS).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial			
CC	inner membrane.			
CC	-1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,			
CC	TO YEAST MRS4; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER			
CC	PROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

34	67	13.3	308	1	UCP3_MOUSE	P56501 mus musculu
35	67	13.3	335	-1	YEA6_YEAST	P39953 saccharomyc
36	67	13.3	373	1	YIA6_YEAST	P40556 saccharomyc
37	67	13.3	419	1	P47B_CANBO	Q00319 candida boi
38	66.5	13.2	292	1	ORT1_YEAST	Q12375 saccharomyc
39	66.5	13.2	587	1	CMC3_CAEEL	Q19529 caenorhabdi
40	66.5	13.2	588	1	CMC2_CAEEL	Q20799 caenorhabdi
41	66	13.1	298	1	ADT2_HUMAN	P03141 homo sapien
42	66	13.1	298	1	ADT3_BOVIN	P32007 bos taurus
43	66	13.1	311	1	TXTP_BOVIN	P79110 bos taurus
44	66	13.1	331	1	ADTL_WHEAT	Q41629 triticum ae
45	66	13.1	331	1	ADT2_WHEAT	Q41630 triticum ae

ALIGNMENTS


```

CC membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC oxoglutarate, adipate, glutarate, 2-oxopimelate, oxaloacetate,
CC citrate and malate. The main physiological role is probably to
CC supply 2-oxoadipate and 2-oxoglutarate from the mitochondrial
CC matrix to the cytosol where they are used in the biosynthesis of
CC lysine and glutamate, respectively, and in lysine catabolism.
CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J75130; CAA99440.1; -.
CC DR EMBL; X92441; CAA63185.1; -.
CC DR SGD; S0005748; ODC2.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 3.
CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
CC FT TRANSMEM 10 30 POTENTIAL.
CC FT TRANSMEM 122 142 POTENTIAL.
CC FT TRANSMEM 215 235 POTENTIAL.
CC FT TRANSMEM 280 300 POTENTIAL.
CC SQ SEQUENCE 307 AA; 34007 MW; 40B9082A64DBA97C CRC64;

Query Match 18.9%; Score 95; DB 1; Length 307;
Best Local Similarity 27.7%; Pred. No. 0.0015;
Matches 23; Conservative 17; Mismatches 43; Indels 0; Gaps 0;

QY 11 RYRNVLALWRIIRTEGLWRPMRGVNTATGAGPAHALYFACYEKLKLTSLDVHPGGNS 70
DB 58 RYGVIDLUKKIVKGGFRLRGISSPMLMEAPKATKFCADNQYQKIFKLNFTNETT 117
QY 71 HTAAGAAGCVATILHDAANPAE 93
DB 118 QKISIAAGASAGTAAVIVPFE 140

RESULT 4
BT1_MAIZE
ID BT1_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN BT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Ilingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT "Analysis of maize brittle-1 alleles and a defective
RT suppressor-mutator-induced mutable allele.";
RL Plant Cell 3:1337-1348(1991).
CC -!- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -!- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -!- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.

```

```

CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M79333; AAA33438.1; -.
CC DR PIR; JQ1459; JQ1459.
CC DR MaizeDB; 47578; -.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 1.
CC KW Transit peptide; Chloroplast; Amyloplast; Transmembrane.
CC FT TRANSIT 1 75 CHLOROPLAST (POTENTIAL).
CC FT CHAIN 76 436 BRITTLE-1 PROTEIN.
CC FT TRANSMEM 229 247 POTENTIAL.
CC FT TRANSMEM 327 347 POTENTIAL.
CC SQ SEQUENCE 436 AA; 46627 MW; 9600C05F603E9DAE CRC64;

Query Match 18.5%; Score 93; DB 1; Length 436;
Best Local Similarity 41.2%; Pred. No. 0.0038;
Matches 21; Conservative 7; Mismatches 23; Indels 0; Gaps 0;

QY 12 YRNVLALWRIIRTEGLWRPMRGVNTATGAGPAHALYFACYEKLKLTSLD 62
DB 365 YQNVLHAIYCIILKEGAGGLYRGLGPGSCIKLMPAGIAFCWEACKKILVD 415

RESULT 5
CG69_HUMAN
ID CG69_HUMAN STANDARD; PRT; 359 AA.
AC Q9BZJ4; Q9UF66; Q9Y379; Q9P182;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial carrier protein CGI-69 (PRO2163).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX PubMed=11139402;
RA Yu X.X., Lewin D.A., Zhong A., Brush J., Schow P.W., Sherwood S.W.,
RA Pan G., Adams S.H.;
RT "Overexpression of the human 2-oxoglutarate carrier lowers
RT mitochondrial membrane potential in HEK-293 cells; contrast with the
RT unique cold-induced mitochondrial carrier CGI-69.";
RL Biochem. J. 353:369-375(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,

```


RT and of a tRNA^{Thr} gene.";
 RL Yeast 13:369-372(1997).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (Potential).
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X99228; CAA67613.1; -;
 CC DR EMBL; 273042; CAA97286.1; -;
 CC DR SGD; S0003489; YGR257C.
 CC DR InterPro; IPR001993; Mitoch_carrier.
 CC DR Pfam; PF00153; mito_carr; 3.
 CC DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 CC KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
 CC KW Transmembrane; Transport.
 CC FT TRANSMEM 17 36 POTENTIAL.
 CC FT TRANSMEM 124 140 POTENTIAL.
 CC FT TRANSMEM 162 182 POTENTIAL.
 CC FT TRANSMEM 268 286 POTENTIAL.
 CC SQ SEQUENCE 366 AA; 40763 MW; B0358B6EBE818CB1E CRC64;
 CC -----
 CC Query Match 15.5%; Score 78; DB 1; Length 366;
 CC Best Local Similarity 34.0%; Pred. NO. 0.16;
 CC Matches 16; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
 CC -----
 CC QY 11 RYRNVEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKIK 57
 CC DB 101 RFNGTLEAFTKIASVEGITSRLMRGSLTLMLAIPANMVFSGYEYIR 147
 CC -----
 CC RESULT 11
 CC ADT_ANOGA
 CC ID ADT_ANOGA STANDARD; PRT; 301 AA.
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
 CC DE ADP,ATP carrier protein (ADP/ATP translocase) (Adenine nucleotide
 CC DE translocator) (ANT).
 CC OS Anopheles gambiae (African malaria mosquito).
 CC OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 CC OC Anopheles.
 CC OX NCBI_TaxID=7165;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=G3;
 CC RX MEDLINE=943348635; PubMed=8069414;
 CC RA Beard C.B., Crews-Oyen A.E., Kumar V.K., Collins F.H.:
 CC FT "A CDNA encoding an ADP/ATP carrier from the mosquito Anopheles
 CC FT gambiae.";
 CC RL Insect Mol. Biol. 3:35-40(1994).
 CC CC -!- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
 CC CC MITOCHONDRIAL INNER MEMBRANE.
 CC CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC CC inner membrane.
 CC CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; L11618; AAB04104.1; -;
 CC DR EMBL; L11617; AAB04105.1; -;
 CC DR InterPro; IPR002067; Mit_carrier.
 CC DR InterPro; IPR001993; Mitoch_carrier.
 CC DR Pfam; PF00153; mito_carr; 3.
 CC DR PRINTS; PR00926; MITOCARRIER.
 CC DR PROSITE; PS00215; MITOCH_CARRIER; 3.
 CC KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 CC FT TRANSMEM 14 31 1 (POTENTIAL).
 CC FT TRANSMEM 75 93 2 (POTENTIAL).
 CC FT TRANSMEM 119 136 3 (POTENTIAL).
 CC FT TRANSMEM 178 197 4 (POTENTIAL).
 CC FT TRANSMEM 216 233 5 (POTENTIAL).
 CC FT TRANSMEM 275 293 6 (POTENTIAL).
 CC SQ SEQUENCE 301 AA; 32863 MW; 4CC9E17C9F8DA08B CRC64;
 CC -----
 CC Query Match 15.4%; Score 77.5; DB 1; Length 301;
 CC Best Local Similarity 28.2%; Pred. NO. 0.15;
 CC Matches 20; Conservative 13; Mismatches 35; Indels 3; Gaps 1;
 CC -----
 CC QY 12 YRNVEALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKIKLSDVIHFGGNSH 71
 CC DB 156 FNGLLDCLKTKTKSGDIIGLYRGVNVQGIITYRAAYFGCFDTAKGMLPD---PKNTSI 212
 CC -----
 CC QY 72 IANGAGCVAT 82
 CC DB 213 FVSWAIAQVVT 223
 CC -----
 CC RESULT 12
 CC MZOM_MOUSE
 CC ID MZOM_MOUSE STANDARD; PRT; 313 AA.
 CC AC Q9CR62;
 CC DT 01-MAR-2002 (Rel. 41, Created)
 CC DT 01-MAR-2002 (Rel. 41, Last sequence update)
 CC DT 01-MAR-2002 (Rel. 41, Last annotation update)
 CC DE Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP).
 CC GN SLC25A11.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=C57BL/6J; TISSUE=Tongue;
 CC RX MEDLINE=21085660; PubMed=11217851;
 CC RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 CC RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 CC RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 CC RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 CC RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 CC RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 CC RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 CC RA Schriml L.W., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 CC RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 CC RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 CC RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 CC RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 CC RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 CC RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 CC RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 CC RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 CC RA Wushaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 CC RA Hayashizaki Y.;
 CC FT "Functional annotation of a full-length mouse cDNA collection.";
 CC RL Nature 409:685-690(2001).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.

Result No.	Score	Query Match	Length	DB	ID	Description
1	496	98.8	364	4	Q96A46	Q96A46 homo sapien
2	351	69.9	347	4	Q9NY22	Q9NY22 homo sapien
3	348	69.3	338	11	Q92OG8	Q92OG8 mus musculus
4	280	55.8	155	4	Q969S1	Q969S1 homo sapien
5	276	55.0	182	11	Q91XR0	Q91XR0 mus musculus
6	216.5	43.1	312	5	Q23125	Q23125 caenorhabdi
7	171.5	34.2	379	5	Q9VAV3	Q9VAV3 drosophila
8	168	33.5	303	5	Q94638	Q94638 onchocerca
9	167.5	33.4	380	5	Q9NHV6	Q9NHV6 drosophila
10	166	33.1	301	5	Q94634	Q94634 onchocerca
11	135	26.9	781	10	Q9LMJ6	Q9LMJ6 arabidopsis
12	130.5	26.0	289	10	Q82049	Q82049 ribes nigru
13	130	25.9	331	10	Q64731	Q64731 arabidopsis
14	118	23.5	303	3	Q14281	Q14281 schizosacch
15	115	22.9	320	11	Q921P8	Q921P8 mus musculus
16	114	22.7	345	11	Q9CVJ1	Q9CVJ1 mus musculus

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 160 SDVIHPGGNSHIANGAAGCVATLLHDAAMNPAE 192

```
Query Match      98.8%; Score 496; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2e-47;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 2

Q9NYZ2 ID Q9NYZ2 PRELIMINARY; PRT; 347 AA.
 AC Q9NYZ2
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HT015 PROTEIN.
 GN HT015.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=HYPOPHALAMUS;
 RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
 RT "A novel gene expressed in human hypothalamus.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF223466; AAF64141.1; -
 DR InterPro; IPR001993; Mitoch_carrier.
 DR Pfam; PF00153; mito_carr.3
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN.2.
 SQ SEQUENCE 347 AA; 37828 MW; 700DE61B230E001E CRC64;

Query Match 69.9%; Score 351; DB 4; Length 347;
 Best Local Similarity 67.7%; Pred. No. 3.3e-31;
 Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
 Db 73 MOSLSPDPKQYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMYFACYENMKRTL 132
 QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMPAPAE 93
 Db 133 NDVFHOGNSHLANGIAGSMATLLHDAVMNPAPAE 165

RESULT 3

Q920G8 ID Q920G8 PRELIMINARY; PRT; 338 AA.
 AC Q920G8
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MITOCHONDRIAL SOLUTE CARRIER-LIKE PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Li Q.-Z., Ruan Q.-G., Eckenrode S., Shi J.-D., Cruz P., Wang C.-Y.,
 RT She J.-X.;
 RL "A new gene which is highly expressed in NOD mice spleen.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF288621; AAL23859.1; -
 SQ SEQUENCE 338 AA; 37510 MW; BB35B1F70C56A3FE CRC64;

Query Match 69.3%; Score 348; DB 11; Length 338;
 Best Local Similarity 68.8%; Pred. No. 7e-31;
 Matches 64; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
 Db 73 MOSLNPDPKARYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMYFACYENMKRTL 132
 QY 61 SDVIHPGNSHIANGAGCVATLLHDAAMPAPAE 93
 Db 133 NDVFHOGNSHLANGIAGSMATLLHDAVMNPAPAE 165

RESULT 4

Q969S1 ID Q969S1 PRELIMINARY; PRT; 155 AA.
 AC Q969S1
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MITOCHONDRIA SOLUTE CARRIER PROTEIN (HYPOTHETICAL 16.8 KDA
 DE PROTEIN).
 GN MSCP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Li Q., Eckenrode S., Ruan Q., Wang C., Shi J., McIndoe R.A., She J.;
 RT "Molecular cloning of a novel mitochondrial solute carrier protein
 RT (MSCP) gene from mouse and human and its down-regulation in mouse
 RT spleen during the maturation of the immune system.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS, AND LEIOMYOSARCOMA;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032628; AAK38154.1; -
 DR EMBL; BC015013; AAH15013.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 155 AA; 16832 MW; 8ACB98A483C8E6EF CRC64;

Query Match 55.8%; Score 280; DB 4; Length 155;
 Best Local Similarity 65.3%; Pred. No. 1.2e-23;
 Matches 49; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MOSLQPPDPAARYNRVLEALWRIIRTEGLWRPMTGLNTATGAGPAHALYFACYEKLKRTL 60
 Db 73 MOSLSPDPKQYTSIYGALKKIMRTGFWRLRGVNVIMGAGPAHAMYFACYENMKRTL 132
 QY 61 SDVIHPGNSHIANG 75
 Db 133 NDVFHOGNSHLANG 147

RESULT 5

Q91ZY0 ID Q91ZY0 PRELIMINARY; PRT; 182 AA.
 AC Q91ZY0
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MITOCHONDRIAL CARRIER-LIKE PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Li Q., Eckenrode S., Wang C., Ruan Q., Shi J., McIndoe R.A., She J.;
 RT "A novel mouse mitochondrial carrier protein gene is up-regulated from
 RT young to adult NOD mice.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF361699; AAL27990.1; -
 SQ SEQUENCE 182 AA; 19838 MW; 18E2C5E801228693 CRC64;

Query Match 55.0%; Score 276; DB 11; Length 182;
 Best Local Similarity 66.7%; Pred. No. 4e-23;
 Matches 50; Conservative 9; Mismatches 16; Indels 0; Gaps 0;


```

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| ||| :| |||:||||| |||||:||||| :|:|
DB 73 MOSLNPDPKARTSYIGALKRIMHTGEFWRPLRGLNVMGAGPAHAMFYFACYENKRTL 132
    :|| |||||:||||
QY 61 SDVHPHGGNSHIANG 75
    :|| |||||:||||
DB 133 NDVESHGNSHLANG 147
    :|| |||||:||||

RESULT 6
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE W02B12.9 PROTEIN.
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=990613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z66521; CA91399.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 43.1%; Score 216.5; DB 5; Length 312;
Best Local Similarity 50.0%; Pred. No. 3.5e-16;
Matches 47; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| :| |||:||||| |||||:||||| :|:|
DB 45 MOSLCPCPETKCPPTVHSLMSIVKREGWLRPLRGYNVAAGSMPAHALYFTVYERKMGYL 104
    :| :| :| |||:||||| |||||:||||| :|:|
QY 61 SDVHPHGGNSH-IANGAAGCVATLLHDAAMNPAE 93
    :| :| :| |||:||||| |||||:||||| :|:|
DB 105 TG--NSAGHSNTLAYGASGVVATLLHDAIMNPAE 136
    :| :| :| |||:||||| |||||:||||| :|:|

RESULT 7
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CG4963 PROTEIN (CH09840P).
GN CG4963.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
RP SEQUENCE FROM N.A.
RC STRAIN=f, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -!- INNER MEMBRANE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AE003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carr; 3.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

Query Match 34.2%; Score 171.5; DB 5; Length 379;
Best Local Similarity 40.9%; Pred. No. 4.8e-11;
Matches 38; Conservative 16; Mismatches 34; Indels 5; Gaps 2;

QY 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
    ||||| ||| :| |||:||||| |||||:||||| :|:|
DB 42 MQSL--SPPTKNMNVSTLRMTREGLLRPIRGASAVVLGAGPAHSLYFAAYENKELT 99
    :|| :| :| |||:||||| |||||:||||| :|:|
QY 61 SDVHPHGGNSHIANGAAGCVATLLHDAAMNPAE 93
    :| :| :| |||:||||| |||||:||||| :|:|
DB 100 AKFTSVRNINVISGA---VATLIHDAISPTD 129
    :| :| :| |||:||||| |||||:||||| :|:|

```

```

RESULT 8
Q94638 ID Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL mitochondrial solute carriers."
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA347E5D617C CRC64;

Query Match 33.5%; Score 168; DB 5; Length 303;
Best Local Similarity 36.6%; Pred. No. 9.2e-11;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

QY 1 MQSLQPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 LQSLPCPCPTSCPTAMHSLMSVMKREGLRLSLKGVNAVLTIPAHAFYTYVYENS KAYL 101
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMNPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 LN--NPRVSNMSYATSGALATVIHDAVMNPAE 132
: : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
Q9NHV6 ID Q9NHV6 PRELIMINARY; PRT; 380 AA.
AC Q9NHV6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOPHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RT "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 33.4%; Score 167.5; DB 5; Length 380;
Best Local Similarity 39.8%; Pred. No. 1.4e-10;

```

```

Matches 37; Conservative 16; Mismatches 35; Indels 5; Gaps 2;

QY 1 MQSLQPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 MQSL--SPPTKNNIVSTLRTMITREGLLRPIRGASAVVLGAGPTHSLSYFAAVEMTKELT 99
: : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMNPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 AKETSVRNLNVIISGA---VATLIHDAISSPTD 129
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q94634 ID Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RT "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL mitochondrial solute carriers."
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45997; AAB19036.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; C0BA8D819FB8EA79 CRC64;

Query Match 33.1%; Score 166; DB 5; Length 301;
Best Local Similarity 36.6%; Pred. No. 1.5e-10;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

QY 1 MQSLQPDPAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 LQSLPCPCPTSCPTAMHSLMSVMKREGLRLSLKGVNAVLTIPAHAFYTYVYENS KAYL 101
:||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 SDVIHPCGNSHIANGAAGCVATLLHDAAMNPAE 93
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 LN--NPRVSNMSYATSGALATVIHDAVMNPAE 132
: : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
Q9LMJ6 ID Q9LMJ6 PRELIMINARY; PRT; 781 AA.
AC Q9LMJ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";

```



```
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; Z99168; CAB16300.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
SQ SEQUENCE 303 AA; 32652 MW; 69975CDE18107AB5 CRC64;

Query Match 23.5%; Score 118; DB 3; Length 303;
Best Local Similarity 33.3%; Pred. No. 3.7e-05;
Matches 26; Conservative 17; Mismatches 31; Indels 4; Gaps 1;

QY 14 NVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKKTLSDVHPGNSHIA 73
DB 60 NTVNSVIKISSTEGYISLWGISSVMGAGPSHAIFYSVLEFFKSK---INASPDRLA 115
QY 74 NGAAGCVATLLHDAAMNP 91
DB 116 SALAGACAITISDAFWTP 133

RESULT 15
Q921P8 PRELIMINARY; PRT; 320 AA.
AC Q921P8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO RIKEN CDNA 5730438N18 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011293; AAH11293.1; -.
SQ SEQUENCE 320 AA; 35050 MW; F40B5912706583CD CRC64;

Query Match 22.9%; Score 115; DB 11; Length 320;
Best Local Similarity 32.9%; Pred. No. 8.5e-05;
Matches 28; Conservative 12; Mismatches 41; Indels 4; Gaps 2;

QY 8 PAARYNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKKTLSDVHPG 67
DB 67 PTVTPGLLQVLKLSILEKEGPKSLFRGLGNLVGVAPSRVAFACYSKAQFNGVFPN 126
QY 68 GNS-HIANGACGCVATLLHDAAMNP 91
DB 127 SNTVHLSAGSAAFVT---NTLMNP 148
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:56:00 ; Search time 138.54 Seconds
(without alignments)
105.029 Million cell updates/sec

Title: US-09-870-113-12
Perfect score: 695
Sequence: 1 MQSLQPDPAARYNRVLEALW.....QFAESTSVLGVNSVTIFYH 131

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	505	72.7	289	21 AAB50383	Human uncoupling p
2	505	72.7	289	22 AAB60113	Human transport pr
3	505	72.7	318	22 AAM41505	Human polypeptide
4	476	68.5	268	22 AAM39719	Human polypeptide
5	352	50.6	299	22 AAM79039	Human protein SQ
6	351	50.5	272	21 AAB42980	Human ORFX ORF2744
7	351	50.5	331	22 AAB60658	Human mitochondria
8	351	50.5	366	22 AAM00938	Human bone marrow
9	322	46.3	677	22 AAU29748	Novel human secret
10	276	39.7	155	21 AAB42966	Human ORFX ORF2730
11	178.5	25.7	379	22 AAB61130	Drosophila melanog

12	165.5	23.8	59	22	AAM85891	Human immune/haema
13	141.5	20.4	289	18	AAW17054	Blackcurrant RIB7
14	139	20.0	278	21	AAG22079	Arabidopsis thalia
15	139	20.0	278	21	AAG43094	Arabidopsis thalia
16	139	20.0	290	21	AAG22078	Arabidopsis thalia
17	139	20.0	290	21	AAG43093	Arabidopsis thalia
18	139	20.0	331	21	AAG22077	Arabidopsis thalia
19	139	20.0	331	21	AAG43092	Arabidopsis thalia
20	121	17.4	130	22	AAM00825	Human bone marrow
21	99.5	14.3	413	22	ABE69108	Drosophila melanog
22	97	14.0	241	21	AAG07172	Arabidopsis thalia
23	97	14.0	289	21	AAG07171	Arabidopsis thalia
24	97	14.0	296	21	AAG07170	Arabidopsis thalia
25	95.5	13.7	310	21	AAG93001	Human protein sequ
26	95.5	13.7	311	21	AAB50382	Human uncoupling p
27	95.5	13.7	311	22	AAM39173	Human polypeptide
28	95.5	13.7	374	22	AAM40959	Human polypeptide
29	95.5	13.7	374	22	AAM40960	Human polypeptide
30	95	13.7	308	21	AAB25780	Human secreted pro
31	95	13.7	308	22	AAB80951	Human carnitine ca
32	95	13.7	308	22	AAB75366	Human secreted pro
33	94.5	13.6	237	21	AAG19645	Arabidopsis thalia
34	94.5	13.6	237	21	AAG50519	Arabidopsis thalia
35	94.5	13.6	284	21	AAG19644	Arabidopsis thalia
36	94.5	13.6	284	21	AAG50518	Arabidopsis thalia
37	94.5	13.6	313	21	AAG19643	Arabidopsis thalia
38	94.5	13.6	313	21	AAG50517	Arabidopsis thalia
39	94.5	13.6	351	21	AAB50378	Human uncoupling p
40	94.5	13.6	365	21	AAB42783	Human ORFX ORF2547
41	92	13.2	96	21	AAG27453	Arabidopsis thalia
42	92	13.2	121	21	AAG27452	Arabidopsis thalia
43	90	12.9	243	21	AAG27858	Arabidopsis thalia
44	90	12.9	285	21	AAG27857	Arabidopsis thalia
45	90	12.9	308	22	AAM38891	Human polypeptide

ALIGNMENTS

RESULT 1

AAB50383

ID AAB50383 standard; Protein; 289 AA.

XX

AC AAB50383;

XX

DT 12-MAR-2001 (first entry)

XX

DE Human uncoupling protein #6.

XX

KW Human; uncoupling protein; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cardiant; vasotropic;

KW cerebroprotective; neuroprotective; antibacterial; ophthalmological;

KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

KW infertility.

XX

OS Homo sapiens.

XX

PN WO200061614-A2.

XX

PD 19-OCT-2000.

XX

PF 06-APR-2000; 2000WO-US09534.

XX

PR 09-APR-1999; 99US-0128701.

PR 08-JUL-1999; 99US-0142821.

PR 18-AUG-1999; 99US-0149448.

PR 12-NOV-1999; 99US-0164751.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

XX

DR WPI; 2000-656322/63.
XX N-PSDB; AAC90457.
XX
PT Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
PS Claim 11; Page 323-324; 343pp; English.
XX
XX The present sequence is a human uncoupling protein. The nucleotide
CC sequences encoding the uncoupling proteins may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
SQ Sequence 289 AA;

Query Match 72.7%; Score 505; DB 21; Length 289;
Best Local Similarity 87.5%; Pred. No. 5.5e-51;
Matches 98; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 MOSLPDPAARYNRVLEALWRIIRTEGLWRPMPRLNVTATGAGPAHALYFACYEKIKKTL 60
DB 25 mqsldpdaarynrvealwriirteglwrpmprglnvtatgagpahalyfacyekikktl 84

QY 61 SDVIHGGNSHIANGAGCVATLLHDAAMNPAE-----GNDSSTYHSVGSC 106
DB 85 sdvihpgnshiangagcvatllhdaamnpaevvkgrmqmynspyrhrvtcd 136

RESULT 2
AAB60113
ID AAB60113 standard; Protein; 289 AA.
XX
XX AAB60113;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33.
XX
XX Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer.
XX
OS Homo sapiens.
XX
XX WO200078953-A2.
XX
XX 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
XX 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI; 2001-04124/05.
DR N-PSDB; AAF27733.
XX
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -

XX Claim 2; Page 133-134; 165pp; English.
XX
XX The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 289 AA;

Query Match 72.7%; Score 505; DB 22; Length 289;
Best Local Similarity 87.5%; Pred. No. 5.5e-51;
Matches 98; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 MOSLPDPAARYNRVLEALWRIIRTEGLWRPMPRLNVTATGAGPAHALYFACYEKIKKTL 60
DB 25 mqsldpdaarynrvealwriirteglwrpmprglnvtatgagpahalyfacyekikktl 84

QY 61 SDVIHGGNSHIANGAGCVATLLHDAAMNPAE-----GNDSSTYHSVGSC 106
DB 85 sdvihpgnshiangagcvatllhdaamnpaevvkgrmqmynspyrhrvtcd 136

RESULT 3
AAM41505
ID AAM41505 standard; Protein; 318 AA.
XX
XX AAM41505;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6436.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60661.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6436; 10078pp; English.
XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 318 AA;

Query Match 72.7%; Score 505; DB 22; Length 318;
 Best Local Similarity 87.5%; Pred. No. 6.2e-51;
 Matches 98; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKKTLL 60
 Db 54 mqsldpdpaaaryrnvlalwriirteglwrpmlgnvtatgagpahalyfacyeklkktl 113

QY 61 SDVIHPGGNSHIANGAGCVATILHDAAMNPAE-----GNDSTSYHVGSC 106

Db 114 sdvihpggnshiangaagcvatilhdaampaevvkqrmqynspyhrvtcdc 165

RESULT 4

AAM39719

ID AAM39719 standard; Protein; 268 AA.

XX AAM39719;

XX 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2864.

KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.
 XX N-PSDB; AAI58875.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

XX Example 4; SEQ ID NO 2864; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 268 AA;

Query Match 68.5%; Score 476; DB 22; Length 268;
 Best Local Similarity 83.0%; Pred. No. 1.3e-47;
 Matches 93; Conservative 1; Mismatches 12; Indels 6; Gaps 1;

QY 1 MQSLQDPDPAARYRNVLALWRIIRTEGLWRPRLNVTATGAGPAHALYFACYEKLKKTLL 60

Db 4 mhsldpdpagryrnvlalwriirteglwrpmlgnvtatgagpahalyfacyeklkktl 63

QY 61 SDVIHPGGNSHIANGAGCVATILHDAAMNPAE-----GNDSTSYHVGSC 106

Db 64 sdvihpggnshiangaagcvatilhdaampaevvkqrmqynspyhrvtcdc 115

RESULT 5

AAM79039

ID AAM79039 standard; Protein; 299 AA.

XX AAM79039;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1701.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52172.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4043; 622lpp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 299 AA;

XX Query Match 50.6%; Score 352; DB 22; Length 299;
XX Best Local Similarity 68.8%; Pred. No. 6.1e-33;
XX Matches 64; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQDPDAARYNNVLEALWRIIRTEGLWRPMRGUNVNTATGAGPAHALYFACYEKLKRTL 60
Db 100 mqlspdpkaqytsygalckimrtgfwprlgvnmimgagpahamyfacyenmkrtl 159

QY 61 SDVHPGNSHANGAACVATLLHDAAMPAPAE 93
Db 160 ndvfhhqgnshlangiagsmatllhdavmpapae 192

RESULT 6

AAAB42980
ID AAB42980 standard; Protein; 272 AA.

XX AAB42980;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF2744 polypeptide sequence SEQ ID NO:5488.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipariatic; antiparkinsonian; neutropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

PN

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAK71789.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 4662-4663; 5507pp; English.

XX AAK74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipariatic; antiparkinsonian; neutropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 272 AA;

XX Query Match 50.5%; Score 351; DB 21; Length 272;

XX Best Local Similarity 67.7%; Pred. No. 7.1e-33;

XX Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLQDPDAARYNNVLEALWRIIRTEGLWRPMRGUNVNTATGAGPAHALYFACYEKLKRTL 60

Db 7 mqlspdpkaqytsygalckimrtgfwprlgvnmimgagpahamyfacyenmkrtl 66

QY 61 SDVHPGNSHANGAACVATLLHDAAMPAPAE 93

Db 67 ndvfhhqgnshlangiagsmatllhdavmpapae 99

RESULT 7

AAAB60658

ID AAB60658 standard; Protein; 331 AA.

XX AAB60658;

XX 04-MAY-2001 (first entry)

XX Human mitochondrial solute carrier protein hMSC-o.

XX Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;

XX preparation; detection.

XX

XX

XX

XX

XX

XX


```

XX OS Homo sapiens.
XX CN1269409-A.
XX PD 11-OCT-2000.
XX PF 17-MAR-2000; 2000CN-0114958.
XX PR 17-MAR-2000; 2000CN-0114958.
XX PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX PI Zhang X, Gao X, Xiao H;
XX WPI; 2001-050544/07.
XX N-PSDB; AAF59920.
XX PT New human mitochondrion solute carrier protein and its nucleic acid -
XX Claim 4; Page 20; 21pp; Chinese.
XX CC The invention relates to a novel human mitochondrial solute carrier
XX protein, hMSC-o (AAF6058), and cDNA encoding it (AAF59920). hMSC-o is
XX expressed in normal human hypothalamus tissue. The invention also relates
XX to the preparation of hMSC-o proteins and nucleic acids, and the
XX detection of hMSC-o proteins and nucleic acids in a sample. The present
XX sequence represents hMSC-o.
XX SQ Sequence 331 AA;

Query Match 50.5%; Score 351; DB 22; Length 331;
Best Local Similarity 67.7%; Pred. No. 9.2e-33;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLOPDAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db mqlslpdkagqtslygalkkkmrtgfwprlgvnmimgagpahamyfacyemkrtl 116

QY 61 SDVIHPGNSHIANGACGCVATLLHDAAMNPAE 93
Db ndvfhhqgnsliangdiagsmatllhdavmpae 149

RESULT 8
AAM00938
ID AAM00938 standard; Protein; 366 AA.
XX AC AAM00938;
XX DT 01-OCT-2001 (first entry)
XX DE Human bone marrow protein, SEQ ID NO: 414.
XX KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; infection; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX OS Homo sapiens.
XX WO200153453-A2.
XX PN 26-JUL-2001.
XX PD 23-DEC-2000; 2000WO-US34960.
XX PF 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.

```

```

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Drmanac RT;
XX WPI; 2001-488707/53.
XX N-PSDB; AAF90057.
XX DR Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX Claim 10; Page 504-505; 648pp; English.
XX CC The present sequence is one of 251 novel human polypeptides encoded
XX by a bone marrow-expressed polynucleotide. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
XX infection, or may result from an autoimmune disorder, a coagulation
XX disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX suppression of an inflammatory response or treatment of a nervous
XX system disorder such as Alzheimer's disease. Detection of the presence
XX or increased expression of the polynucleotide or the protein it
XX encodes is useful for the diagnosis and/or prognosis of one
XX or more types of cancer. The polynucleotide and polypeptide can be
XX used as nutritional sources or supplements and in the screening of
XX chemical compounds as potential drugs.
XX SQ Sequence 366 AA;

Query Match 50.5%; Score 351; DB 22; Length 366;
Best Local Similarity 67.7%; Pred. No. 1.1e-32;
Matches 63; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY 1 MOSLOPDAARYRNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKRTL 60
Db mqlslpdkagqtslygalkkkmrtgfwprlgvnmimgagpahamyfacyemkrtl 160

QY 61 SDVIHPGNSHIANGACGCVATLLHDAAMNPAE 93
Db ndvfhhqgnsliangdiagsmatllhdavmpae 193

RESULT 9
AAU29748
ID AAU29748 standard; Protein; 677 AA.
XX AC AAU29748;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #239.
XX KW Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX WO200179449-A2.
XX PN 25-OCT-2001.
XX PD 16-APR-2001; 2001WO-US08656.
XX PF 18-APR-2000; 2000US-0552929.

```



```
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254057.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX N-PSDB; AAK58672.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254057.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX N-PSDB; AAK58672.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Claim 11; SEQ ID NO 13484; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX Sequence 59 AA;
XX
XX Query Match 23.8%; Score 165.5; DB 22; Length 59;
XX Best Local Similarity 59.3%; Pred. No. 6.6e-12;
XX Matches 35; Conservative 1; Mismatches 12; Indels 11; Gaps 1;
XX
XX QY 8 PAARYNVLEALWRIIRTEGLWRPMPGLNVTATGAGPAHALYFACYEKLKKTLSVHP 66
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 12 pldykn-----ggpmeamrglnvtatgagpxhnglyfacteklkalsdvlhp 59
XX
XX RESULT 13
XX AAW17054
XX ID AAW17054 standard; Protein; 289 AA.
XX AC AAW17054;
XX XX 05-AUG-1997 (first entry)
XX DT Blackcurrant RIB7 polypeptide.
XX DE Blackcurrant; fruit-specific promoter; RIB7; transgenic plant.
XX KW Ribes nigrum strain Ben Alder.
XX OS WO9717452-A1.
XX PN 15-MAY-1997.
XX PD
XX PF 04-NOV-1996; 96WO-EP04807.
XX PR 03-NOV-1995; 95GB-0022558.
XX XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX XX Brennan RM, Taylor MA, Woodhead MR;
XX WPI; 1997-281041/25.
XX N-PSDB; AAT68957.
XX New isolated promoters from blackcurrant fruit - used for driving
XX fruit-specific expression of DNA sequences in transgenic
XX blackcurrant and other non-climacteric fruit
XX Example 4; Page 35-36; 66pp; English.
XX RIB polypeptides (AAW17050-54) are encoded by cDNA clones (AAT68953-
```

CC 57) that exhibit differential expression in blackcurrant fruit
CC during the ripening period of fruit development. Rib7 shows
CC 62% similarity to yeast MRS4, a yeast mitochondrial RNA splicing
CC protein. Rib7 is expressed almost entirely in fruit. The
CC promoter region (AA168952) of the Rib7 gene can be used as a
xx fruit-specific promoter.
SQ Sequence 289 AA;

Query Match 20.4%; Score 141.5; DB 18; Length 289;
Best Local Similarity 31.2%; Pred. No. 3.8e-08;
Matches 40; Conservative 18; Mismatches 51; Indels 19; Gaps 3;

Qy 9 AARYRNVLRLRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKTLSDVIHPGG 68
Db 32 sadsaglrqalgsilkvepgaglyrigamglgagpahavysvymcketfs---hgdp 88
Qy 69 NSHIANGAAGCVATLLHDAAMNPAE-----GNDSSYTHSVGSCCTISLQFAESTSVLV 122
Db 89 snsgahavsgvfatasdavtpmdvvkqrqlqsspykgvvdv-----rrvlv 138
Qy 123 GNSVTLFY 130
Db 139 eegigafy 146

RESULT 14
AAG22079
ID AAG22079 standard; Protein; 278 AA.
AC AAG22079;
XX
XX
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 24871.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24871.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 18-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0145388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149929.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158223.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 139; DB 21; Length 278;
Best Local Similarity 36.2%; Pred. No. 7e-08;
Matches 42; Conservative 15; Mismatches 41; Indels 18; Gaps 5;

QY 1 MOSLOPDPAARYRNVEALRLIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60
Db 13 mqlalrscp-ikpigrqafsrliktdgpsyalyrgiawmglgagpahavysfyevskkfll 71
QY 61 SDVIHPGG--NSHIANGAGCVATLLHDAAMNPAE-----GNDSTYHSVSGC 106
Db 72 s-----ggpnnsaahaisgvfatissdavftpmdmvkvqlqign--gtykgvwdc 130

RESULT 15
AAG43094
ID AAG43094 standard; Protein; 278 AA.
XX
AC AAG43094;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53822.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
```

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135625.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139450.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      20.0%; Score 139; DB 21; Length 278;
Best Local Similarity 36.2%; Pred. No. 7e-08;
Matches 42; Conservative 15; Mismatches 41; Indels 18; Gaps 5;

QY 1 MQSLQDPDAARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKRTL 60
   ||:| | : :| ||:| | ||: |||||: || || |
Db 13 mqlrscp-ikpigrqafrsiiktgdpsalyrgiawmglgagpahavyfsfyevskkfl 71

QY 61 SDVIHPGG--NSHIANGACGCVATLLHDAAMPAAE-----GNDSTYHVSVC 106
   | || | : :| ||: || | : || || | |
Db 72 s-----ggnpnsaahaisgvfatissdavftpmdmvmkqrlqign--gtykgvwdc 120
```

Search completed: August 27, 2002, 03:56:01
Job time: 4913 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 03:58:53 ; Search time 48.11 Seconds
(without alignments)
66.509 Million cell updates/sec

Title: US-09-870-113-12
Perfect score: 695
Sequence: 1 MQSLQPPAARYNRVLEALW.....QFAEESTSVLGVNSVTLFYH 131

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141.5	20.4	289	US-09-068-140A-10	Sequence 10, Appl
2	141.5	20.4	328	US-09-068-140A-15	Sequence 15, Appl
3	94.5	13.6	351	US-08-933-750C-19	Sequence 19, Appl
4	94.5	13.6	351	US-09-234-613-19	Sequence 19, Appl
5	90	12.9	320	US-08-933-750C-12	Sequence 12, Appl
6	90	12.9	320	US-09-234-613-12	Sequence 12, Appl
7	79	11.4	293	US-08-628-291-4	Sequence 4, Appl
8	79	11.4	293	US-09-128-722-4	Sequence 4, Appl
9	79	11.4	317	US-08-628-291-12	Sequence 12, Appl
10	79	11.4	317	US-09-128-722-12	Sequence 12, Appl
11	71	10.2	1094	US-09-268-347-32	Sequence 32, Appl
12	71	10.2	2584	US-08-936-135-4	Sequence 4, Appl
13	69	9.9	306	PCT-US94-09799-1	Sequence 1, Appl
14	68.5	9.9	311	US-08-775-009-32	Sequence 32, Appl
15	68.5	9.9	312	US-09-142-565-2	Sequence 2, Appl
16	68.5	9.9	658	US-08-409-995-5	Sequence 5, Appl
17	68.5	9.9	658	US-08-685-467-5	Sequence 5, Appl
18	68.5	9.9	658	US-08-913-942-5	Sequence 5, Appl
19	68.5	9.9	1098	US-08-409-995-2	Sequence 2, Appl
20	68.5	9.9	1098	US-08-685-467-2	Sequence 2, Appl
21	68.5	9.9	1098	US-09-377-155-32	Sequence 32, Appl
22	68.5	9.9	1098	US-08-913-942-2	Sequence 2, Appl
23	68.5	9.9	1098	US-09-669-974-32	Sequence 32, Appl
24	68.5	9.9	1098	US-09-268-347-44	Sequence 44, Appl
25	67	9.6	432	US-08-937-466-4	Sequence 4, Appl
26	67	9.6	432	US-09-172-528-4	Sequence 4, Appl
27	67	9.6	432	US-09-318-199-4	Sequence 4, Appl

28	67	9.6	432	4	US-09-503-579-4	Sequence 4, Appl
29	66	9.5	2594	4	US-08-718-388-7	Sequence 7, Appl
30	66	9.5	5405	4	US-08-718-388-9	Sequence 9, Appl
31	65.5	9.4	311	2	US-08-775-009-33	Sequence 33, Appl
32	65	9.4	447	4	US-09-160-119-2	Sequence 2, Appl
33	65	9.4	674	4	US-09-160-119-2	Sequence 2, Appl
34	63.5	9.1	298	3	US-08-961-871-10	Sequence 10, Appl
35	63	9.1	308	2	US-08-937-466-2	Sequence 2, Appl
36	63	9.1	308	2	US-09-172-528-2	Sequence 2, Appl
37	63	9.1	308	3	US-09-318-199-2	Sequence 2, Appl
38	63	9.1	308	4	US-09-503-579-2	Sequence 2, Appl
39	63	9.1	571	2	US-08-892-770-5	Sequence 5, Appl
40	62	8.9	469	4	US-09-188-930-339	Sequence 339, Appl
41	62	8.9	888	4	US-09-268-140-4	Sequence 4, Appl
42	61.5	8.8	119	1	US-08-053-171-14	Sequence 14, Appl
43	61.5	8.8	119	1	US-08-053-171-17	Sequence 17, Appl
44	61.5	8.8	119	3	US-08-815-190A-13	Sequence 13, Appl
45	61.5	8.8	2588	3	US-08-936-135-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-068-140A-10
; Sequence 10, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-10

RESULT 7

ZIF. 14014 1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

QY 51 ACYEKLKKTLSDVIHPPGNSHIANGAAGCVATLLHDAAMNPAEGNDSSSTYHSVGSCTCIS 110

Db 80 AQVETLFRSINGVLLPGGANLTHSGYSRVAKIFFTKAL---ESFDNGDFFPVWG-TCLG 135
QY 111 LQFAESTSVLVGN 124
Db 136 L-----EELSLVSN 145

RESULT 11
US-09-268-347-32
; Sequence 32, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-32

Query Match 10.2%; Score 71; DB 4; Length 1094;
Best Local Similarity 27.9%; Pred. No. 8.7;
Matches 29; Conservative 17; Mismatches 46; Indels 12; Gaps 4;
QY 34 GLNVATGAGPAHALYFACYEKL---KKTLSVDVHPGGNSHIANGAAGCVATLLHDAAM 90
Db 150 GVQVTSSENGKAIYFALAKLDMRTATVSDLTITGGSTTTGSATTPKYNVTSTAGLN 209
QY 91 PAEG----NDSSYHSVGSCTCISLQFAESTSV---LVGNSVT 127
Db 210 FARGATGANGDTVHLTNIAS--TLQDTLLNTGVWSKLDNGKIT 251

RESULT 12
US-08-936-135-4
; Sequence 4, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-936-135-4

Query Match 10.2%; Score 71; DB 3; Length 2584;
Best Local Similarity 29.1%; Pred. No. 29;
Matches 32; Conservative 15; Mismatches 35; Indels 28; Gaps 6;
QY 3 SLOPDP-AARYRVNLEALWRIIR-TEGLWRPMBGLNVNTATGAGPAHALYFACYEKLAKTL 60
Db 1320 ALERPRSRERPHISPTRYRILEASNGLTRPLEGLNVAL--ASPLEGLY-ASPGLLYSIL 1376
QY 61 SDVIHPGG-----NSHIANGAAGCVATLLHDAAM 89
Db 1377 EVALARGGLYVALILEITLEGNGLYGLYSHIS--ARGGLASNLYSVAL 1424

RESULT 13
PCT-US94-09799-1
; Sequence 1, Application PC/TUS9409799
; GENERAL INFORMATION:
; APPLICANT: Kagan, David
; TITLE OF INVENTION: Method and Composition for Weight Reduction
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Watov & Kipnes, P.C.
; STREET: 186 Princeton-Hightstown Rd, PO Box 247
; CITY: Princeton Junction
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08550
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: DOS 6.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09799
; FILING DATE: 29-AUG-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,179
; FILING DATE: 25 May 1994
; APPLICATION NUMBER: 08/114,313
; FILING DATE: 30 August 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kipnes, Allen R.
; REGISTRATION NUMBER: 28,433
; REFERENCE/DOCKET NUMBER: 489.1.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-243-0330
; TELEFAX: 609-275-1010
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: Linear
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL TYPE: Brown adipose
; FEATURE:
; OTHER INFORMATION: Cassard, A.M. et al. publication
; OTHER INFORMATION: Genbank - Locus 37607
; PUBLICATION INFORMATION:
; AUTHORS: Cassard, A.M. et al.
; TITLE: Human Uncoupling Protein Gene: Structure,
PCT-US94-09799-1

C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #taxon
C.Accession: A68205
R.Title: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, K.A.; Lander, E.S.; Chu, P.H.; Holt, T.F.; Nelson, D.E.; Chao, M.; Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conwell, B.L.; Gnanapavan, S.; Hake, R.; Harber, P.B.; Huiziar, L.; Jensen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A.Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Starker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Dubois, J.P.
A.Title: Sequence and analysis of chromosome 1 of the genome of Arabidopsis thaliana.
A.Reference number: A86141; PMID:21016719

Db 72 KNMLSQISHISTEGTLALWKGVQSVILGAGPAHAHVFGTYEFCCKNLIDSSDTQTHHPF 131
QY 73 ANGAACGCVATLLHDAAMNPAEG-----NDSST-----YHVSVCSTC 108
Db 132 KTAISGACATTASDALMNPFDTIKORIQLNSTASVQWTTKQIQVQSEGLAAFYSYPTTLV 191
QY 109 ISLOFA-----EESTSVL 121
Db 192 MNIPFAAFNEFVIESSTKEL 211

RESULT 5
T39149
Probable RNA splicing protein mitochondrial carrier protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39149
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: 221748
A:Accession: T39149
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-303 <OLI>
A:Cross-References: EMBL:Z99168; PIDN:CAB16300.1; GSPDB:GN00066; SPDB:SPAC8C9.12c
A:Experimental source: strain 972h-; cosmid c8C9
C:Genetics:
A:Gene: SPDB:SPAC8C9.12c
A:Map position: 1
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 17.1%; Score 119; DB 2; Length 303;
Best Local Similarity 27.6%; Pred. No. 0.00022;
Matches 34; Conservative 21; Mismatches 48; Indels 20; Gaps 3;
QY 14 NVLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKTLSDVIHPGGNSHIA 73
Db 60 NIVNSVIKISSTEGVYSLRWGISSVIMGAGPSHAIFYSVLEPFKSK-----INASPRPLA 115
QY 74 NGAAGCVATLLHDAAMNPAE-----GNDSSYHVSVCSTCISLQPAESTSVLVGNSVT 127
Db 116 SALAGACAITISDAFPTFDVIKQRMQLPSRKYKSALHC-----ATTVFRNEGLG 165
QY 128 LFY 130
Db 166 AFY 168

RESULT 6
S13533
mRNA splice defect-suppressing mitochondrial carrier MRS4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR052c
C:Species: Saccharomyces cerevisiae
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jul-2000
C:Accession: S13533; S38126
R:Wiesenberger, G.; Link, T.A.; von Absen, U.; Waldherr, M.; Schweyen, R.J.
J. Mol. Biol. 217, 23-37, 1991
A:Title: MRS4 and MRS4, two suppressors of mRNA splicing defects in yeast, are new members of the ADP/ATP carrier family
A:Reference number: S13532; MUID:91108815
A:Accession: S13533
A:Molecule type: DNA
A:Residues: 1-304 <JMO>
A:Cross-References: EMBL:X56444; NID:g3995; PIDN:CAA39828.1; PID:g3996
R:Visser, S.; Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38118
A:Accession: S38126
A:Molecule type: DNA
A:Residues: 1-304 <VIS>
A:Cross-References: EMBL:Z28277; NID:g486506; PID:g486507; MIPS:YKR052c
A:Experimental source: strain S288C
C:Genetics:

A:Gene: SGD:MRS4
A:Cross-References: SGD:S0001760; MIPS:YKR052c
A:Map position: 11R
A:Genome: nuclear
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C:Keywords: duplication; mitochondrion; transmembrane protein
F:20-109/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:117-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:206-301/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 15.3%; Score 106; DB 2; Length 304;
Best Local Similarity 32.5%; Pred. No. 0.0046;
Matches 25; Conservative 12; Mismatches 40; Indels 0; Gaps 0;
QY 15 VLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLKTLSDVIHPGGNSHIAN 74
Db 64 MTSQISKISTMEGSMALWKGVSILGAGPAHAHVFGTYEFCCKARLISPEDMTHQPMKT 123
QY 75 GAAGCVATLLHDAAMNP 91
Db 124 ALSGTATTAAADALMNP 140

RESULT 7
S44092
Probable carrier protein c2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C:Accession: S44092; T24753
R:Runswick, M.J.; Philippides, A.; Lauria, G.; Walker, J.E.
submitted to the EMBL Data Library, November 1993
A:Description: Extension of the mitochondrial transport superfamily: sequences of five different proteins from the same family
A:Reference number: S44090
A:Accession: S44092
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <RUN>
A:Cross-References: EMBL:X76116; NID:g472899; PIDN:CAA53722.1; PID:g472900
R:Lloyd, C.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19932
A:Accession: T24753
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-384 <WIL>
A:Cross-References: EMBL:Z49070; PIDN:CAA88869.1; GSPDB:GN00020; CESP:T09F3.2
A:Experimental source: clone T09F3
C:Genetics:
A:Gene: CESP:T09F3.2
A:Map position: 2
A:Introns: 82/1; 113/1; 153/2; 179/2; 259/3; 359/3
C:Superfamily: Caenorhabditis probable carrier protein c2; ADP,ATP carrier protein repeat homology <ACP1>
F:4-43,141-193/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:45-67/Region: serine-rich
F:201-285/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:296-381/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 14.6%; Score 101.5; DB 1; Length 384;
Best Local Similarity 35.1%; Pred. No. 0.017;
Matches 33; Conservative 17; Mismatches 31; Indels 13; Gaps 5;
QY 5 QPDPAARYRN-VLEALWRIIRTEGLWRPRLGNVTATGAGPAHALYFACYEKLK--TUS 61
Db 137 QPPTAARRGTIVIKYITQIKTEGICALYKGLIPNLVGVAPSKAVFYTYSTSKRFWES 196
QY 62 DVTHPGNSHIAN---GAAGCVATLLHDAAMNP 91
Db 197 EVLIP--NSAIHVMSAGSAGFVAA-----SAVNP 224
RESULT 8

B96830
hypothetical protein F19k16.14 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C2:Mar-02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96830
R:Theologidis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C2:Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: B96830
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <SFO>
A:Cross-references: GB:AE0051173; NID:g6453859; PIDN:AAF09043.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19k16.14
A:Map position: 1

Query Match	13.7%	Score 95;	DB 2;	Length 296;
Best Local Similarity	28.5%;	Pred. No.	0.057;	
Matches	35;	Conservative	17;	Mismatches 43; Indels 28; Gaps 7;
QY	22	IIRTEGLWRPMRGUNVTATGAGPAHALYFCAYEKLKLTSLDVHPG----	GNSH-----I	72
	:	: :	:	:
	:	:	:	:
	:	:	:	:
Db	150	ILRROGLQGLVRLTIVLRDAPAHGLYFWTYEVVRERL----	HPCRKTKQENLRMLV	205
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY	73	ANGAAG-----CVATLLHDAAAMPAGNDSSTHVSGSCTCISLQFAESTSVL----	VG	123
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db	206	AGGLAGVASWVACYPLDWVKTRLQQHG-----AYEGTADCFRKGVK--QEYTVLWRGLG		259

Qy	124	NSV	126
		:	
Dp	260	TAV	262

RESULT 9
 S60949
 probable phosphate transport protein, mitochondrial - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein O5644; protein YOR222w; protein YOR50-12
 C;Species: Saccharomyces cerevisiae
 C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
 C;Accession: S60949; S67115; B37984; S71724

R;Galisson, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A;Reference number: S60938
A;Accession: S60949
A;Molecule type: DNA
A;Residues: 1-307 <GAL>
A;Cross-references: EMBL:X92441; NID:g1050762; PID:g1050774
R;Boyer, J.; Fairhead, C.; Gallon, L.; Galisson, F.; Michaux, G.; Thierry, A.; Dujon, B.
submitted to the Protein Sequence Database, July 1996

A:Reference number: S67104
A:Accession: S67115
A:Molecule type: DNA
A:Residues: 1-307 <BOY>
A:Cross-references: EMBL:275130; NID:g1420513; PID:g1420514; MIPS:YOR222w
A:Experimental source: Strain S288C
R:Guerin, B.; Bukusoglu, C.; Rakotomanana, F.; Wohlrab, H.
J. Biol. Chem. 285, 19736-19741, 1990
A:Title: Mitochondrial phosphate transport. N-ethylmaleimide insensitivity correlates with
A:Reference number: A37984; MUID:91060585
A:Accession: B37984
A:Molecule type: protein
A:Residues: 30,'X',32,'I',34-39 <GUE>

R:Galisson, F.; Dujon, B.
Yeast 12, 877-885, 1996
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chr
A:Reference number: S711713; MUID:96437977
A:Accession: S711724
A:Molecule type: DNA
A:Residues: 1-307 <GAW>
A:Cross-references: EMBL:X92441; NID:g1050762; PIDN:CRA63185.1; PID:g1050774
C:Genetics:
A:Map position: 15R
A:Note: YOR222W
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homol
C:Keywords: duplication; mitochondrion; transmembrane protein
F:9-107/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F:115-201/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F:208-300/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 13.7% Score 95 DB 2 Length 307;
Best Local Similarity 27.7% Pred. No. 0.059;
Matches 23; Conservative 17; Mismatches 43; Indels 0; Gaps 0;

QY 11 RYRNVEALWRIIRTEGLWRPMRGLNVTATGAGPAHALFYACYEKIKKTLSDVIHPGNS 70
|| | : : | : | || : | | | : : | : :
Db 58 RYNGVIDCKIVKEGESFRIYRGISSPMLMEAPKRATFACNDOYOKIEKNLFNNETT 117

Qy	71	HIANGAGCVATLLHDAAMNPAE	93
Db	118	QKISIAAGASAGMTEAAVIVPFE	14

RESULT 10
D84613
hypothetical protein At2g22500 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84613
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis tha
A:Reference number: A84420: MUID:20083487

A;Residues: 1-313 <STU>
A;Cross-references: GB:AF002093; NID:g4544443; PIDN:AAD22351.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g22500
A;Map position: 2
C;Superfamily: ADP.ATP carrier protein; ADP.ATP carrier protein repeat homolog

Query Match	Score	DB 2	Length
Best Local Similarity	13.6%	94.5	313
Matches	31.8%	Pred. No. 0.068	
Conservative	23	Mismatches	28
Indels	27	Gaps	3

Qy 12 YRNVEALWRIIRTEGLWRPMRGLNVTATGAGPAHALFYACYEKIKKLTSD--VHPGGN 69
|::||: ::|| |::| : |::|: : : |
Dd 161 YKSVDATOMIRGEGVTSWRGSSITINRAMLVTSOLASYSVKETILEKGLKDLGL 220

QY 70 SHI-ANGAGCVATLLHDAAMP AE 93
 : : : : : : : : : : : : :
pb 221 THVSASFAGFVASV---ASNPD 241

```

RESULT 11
T43493
hypothetical protein DKFZp434C119.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43493

```


T20229
hypothetical protein C54G10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20229
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20229
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-310 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99811.1; GSPDB:GN00023; CESP:C54G10.4
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.4
A:Map position: 5
A:Introns: 14/1; 63/1; 155/2; 256/2; 299/3
C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 12.88; Score 89; DB 2; Length 310;
Best Local Similarity 25.58; Pred. No. 0.24; 55; Indels 38; Gaps 8;
Matches 39; Conservative 21; Mismatches 55; Indels 38; Gaps 8;
Qy 4 LQPDPA-ARYENVLEALWRIIRTEGLWRPMPRGLNVTATGAGPAHALYFACYEKLK---- 58
| | | | | : : : : : | | | | | : : : : :
Db 125 IQDDKAHTKFNPIDATKQLLRTHGLKSLTRGFLATVARDAPAFGVYFASYENWARSCK 184
| | | | | : : : : : | | | | | : : : : :
Qy 59 -----TSLDVIHPGNGSHIANGAGCVATLLH---DAAMNPAEGNDS--STYHSVGSCTC 108
| | | | | : : : : : | | | | | : : : : :
Db 185 DGETSTLS-----SGQLIFAGGTAGMLSWLNFNYQTDIIKSRFQADNSYKSYM-----C 233
Qy 109 ISLOFAEES-----TSLV-----GNSVTLF 129
| | | | | : : : : : | | | | | : : : : :
Db 234 IKQYILERGYRGFFVGLNSALIRAFPSNAATFF 266

Search completed: August 27, 2002, 03:57:45
Job time: 4951 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:16:49 ; Search time 37.5 Seconds
(without alignments)
135.260 Million cell updates/sec

Title: US-09-870-113-12

Perfect score: 695

Sequence: 1 MSLQPDPAARYNVLALW.....QFAESTSVLGVNSVTLFYH 131

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125.5	18.1	314	1 MRS3_YEAST	P10566 saccharomyc
2	106	15.3	304	1 MRS4_YEAST	P23500 saccharomyc
3	95	13.7	307	1 ODC2_YEAST	Q99297 saccharomyc
4	94.5	13.6	359	1 CG69_HUMAN	Q9bz14 homo sapien
5	93	13.4	436	1 BTL1_MAIZE	P29518 zea mays (m
6	90	12.9	320	1 DNC_HUMAN	Q9hc21 homo sapien
7	89.5	12.9	161	1 PHAB_CYACA	Q9tlis8 cyanidium c
8	86.5	12.4	161	1 PHAB_GALSU	P00319 galdieria s
9	86.5	12.4	322	1 SFCL_YEAST	P33303 saccharomyc
10	82.5	11.9	161	1 PHAB_SYNY3	Q01952 synchocyst
11	82.5	11.9	161	1 PHAB_SYNY4	Q02924 synchocyst
12	80.5	11.6	162	1 PHAB_FREDI	P16571 fremyella d
13	80.5	11.6	313	1 M2OM_BOVIN	P22292 bos taurus
14	79.5	11.4	161	1 PHAB_MASLA	P00318 mastigoclad
15	79.5	11.4	313	1 M2OM_HUMAN	Q02978 homo sapien
16	79.5	11.4	328	1 Y051_CAEEL	Q09461 caenorhabdi
17	79	11.4	317	1 GGH_RAT	Q62867 rattus norv
18	78.5	11.3	161	1 PHAB_CYAPA	P00320 cyanophora
19	78	11.2	366	1 YG5F_YEAST	P53320 saccharomyc
20	77.5	11.2	301	1 ADT_ANOGA	Q27238 anopheles g
21	77.5	11.2	313	1 M2OM_MOUSE	Q9cr62 mus musculu
22	77	11.1	368	1 YM39_YEAST	Q03829 saccharomyc
23	76.5	11.0	161	1 PHAB_SYNEL	P50031 synchococc
24	76.5	11.0	1320	1 PUTA_SALTY	P10503 salmonella
25	76	10.9	315	1 SA18_HUMAN	Q9hik4 homo sapien
26	75.5	10.9	161	1 PHAB_AGLNE	P28556 aglaothamni
27	75.5	10.9	161	1 PHAB_ANASP	P80557 anabaena sp
28	75.5	10.9	161	1 PHAB_ANAVA	P00317 anabaena va
29	75.5	10.9	161	1 PHAB_SYNP6	P06113 synchococc
30	74.5	10.7	161	1 PHAB_SPIPL	P72505 spirulina p
31	74	10.6	309	1 UCP2_PIG	Q97562 sus scrofa
32	73.5	10.6	161	1 PHAB_ANACY	P07326 anabaena cy
33	73.5	10.6	287	1 DIC_HUMAN	Q9ubx3 homo sapien

34	73.5	10.6	287	1 DIC_MOUSE	Q9qzd8 mus musculu
35	73.5	10.6	313	1 M2OM_RAT	P97700 rattus norv
36	73	10.5	312	1 SA18_MOUSE	Q9db41 mus musculu
37	72	10.4	695	1 CMCI_DROME	Q9a731 drosophila
38	71.5	10.3	315	1 MFT_HUMAN	Q9h2d1 homo sapien
39	71	10.2	308	1 UCP3_RAT	P56499 rattus norv
40	70	10.1	299	1 YFBH_SALTY	O52326 salmonella
41	70	10.1	301	1 FPG_RHIME	O59752 rhizobium m
42	70	10.1	308	1 ADT_CHLRE	P27080 chlamydomon
43	69.5	10.0	292	1 ORT1_YEAST	Q12375 saccharomyc
44	69.5	10.0	299	1 TXTP_YEAST	P38152 saccharomyc
45	69.5	10.0	1312	1 PUTA_KLEAE	O52485 klebsiella

ALIGNMENTS

RESULT	1			
MRS3_YEAST				
ID	MRS3_YEAST	STANDARD;	PRT;	314 AA.
AC	P10566;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Mitochondrial RNA splicing protein MRS3.			
GN	MRS3 OR YJ1133W OR J0675.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=IC8/R101;			
RX	MEDLINE=88121698; PubMed=2448588;			
RA	Schmidt C., Soellner T., Schweyen R.J.;			
RT	"Nuclear suppression of a mitochondrial RNA splice defect: nucleotide			
RT	sequence and disruption of the MRS3 gene.";			
RL	Mol. Gen. Genet. 210:145-152(1987).			
RN	[2]			
RP	REVISIONS, SEQUENCE FROM N.A.			
RC	STRAIN=M1301;			
RX	MEDLINE=91108815; PubMed=1703236;			
RA	Wiesenberg G., Link T.A., von Ahsen U., Waldherr M., Schweyen R.J.;			
RT	"MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast,			
RT	are new members of the mitochondrial carrier family.";			
RL	J. Mol. Biol. 217:23-37(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / FY1679;			
RX	MEDLINE=96408771; PubMed=8813765;			
RA	Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;			
RT	"Sequence analysis of a 40.7 kb segment from the left arm of yeast			
RT	chromosome X reveals 14 known genes and 13 new open reading frames			
RT	including homologues of genes clustered on the right arm of			
RT	chromosome XI.";			
RL	Yeast 12:787-797(1996).			
CC	-1- FUNCTION: MRS3 SUPPRESSES A MITOCHONDRIAL SPLICE DEFECT IN THE			
CC	FIRST INTRON OF THE COB GENE. IT MAY ACT AS A CARRIER, EXERTING			
CC	ITS SUPPRESSOR ACTIVITY VIA MODULATION OF SOLUTE CONCENTRATIONS IN			
CC	THE MITOCHONDRIUM (POSSIBLY OF CATIONS).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial			
CC	inner membrane.			
CC	-1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. STRONG,			
CC	TO YEAST MRS4; BOTH PROTEINS MAY BE ISOFORMS OF THE SAME CARRIER			
CC	PROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			


```
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT VARSPLIC 64 72 LPSLSOSTG -> W (IN ISOFORM 2).
FT VARIANT 247 L -> F.
FT CONFLICT 266 266 /FTID=VAR_012756.
FT SEQUENCE 359 AA; 39214 MW; 952AA3DB5F5F9BD1 CRC64;

Query Match 13.6%; Score 94.5; DB 1; Length 359;
Best Local Similarity 29.6%; Pred. No. 0.012;
Matches 34; Conservative 15; Mismatches 39; Indels 27; Gaps 6;

QY 7 DPAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKLT--- 60
DB 100 DP-TRFTGTMDAFYKIVRHEGRTLW---SGLPATLVMTVPATAIYFTAYDQLKFLGCR 155
QY 61 ---SDVTHPGGSHIANGAAGCVATLLHDAAMPAE-----GNDSTYHSVSGC 106
DB 156 ALTSOLIAP-----MWAGALARLGTIVVISPLELMTKQLQAHVSYRELGC 202

RESULT 5
BT1_MAIZE
ID B11_MAIZE STANDARD; PRT; 436 AA.
AC P29518;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brittle-1 protein, chloroplast precursor.
GN Brl.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005685; PubMed=1668652;
RA Sullivan T.D., Strelow L.I., Illingworth C.A., Phillips R.L.,
RA Nelson O.E. Jr.;
RT "Analysis of maize brittle-1 alleles and a defective
RT suppressor-mutator-induced mutable allele.";
RL Plant Cell 3:1337-1348(1991).
CC -1- FUNCTION: COULD PLAY A ROLE IN AMYLOPLAST MEMBRANE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
CC -1- TISSUE SPECIFICITY: ENDOSPERM OF DEVELOPING KERNELS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M79333; AAA33438.1; -.
CC DR PIR; J01459; J01459.
CC DR MaizeDB; 47578; -.
CC DR InterPro; IPR002067; Mit_carrier.
CC DR InterPro; IPR001993; Mitoch_carrier.
CC DR Pfam; PF00153; mito_carr; 3.
CC DR PRINTS; PR00926; MITOCARRIER.
CC DR PROSITE; PS00215; MITOCH_CARRIER; 1.
CC TRANSIT peptide; Chloroplast; Amyloplast; Transmembrane.
FT CHAIN 1 75 CHLOROPLAST (POTENTIAL).
FT TRANSIT 76 436 BRITTLE-1 PROTEIN.
FT TRANSMEM 229 247 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
```

```
SQ SEQUENCE 436 AA; 46627 MW; 9600C05F603E9DAE CRC64;

Query Match 13.4%; Score 93; DB 1; Length 436;
Best Local Similarity 41.2%; Pred. No. 0.021;
Matches 21; Conservative 7; Mismatches 23; Indels 0; Gaps 0;

QY 12 YRNVLALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYEKLKLTSLD 62
DB 365 YQNVLHAIYCLIKKEGAGGLYRGIGPSCIKLMPAAGIAFWCYEACKKILVD 415

RESULT 6
DNC_HUMAN
ID DNC_HUMAN STANDARD; PRT; 320 AA.
AC Q9HC21;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial deoxynucleotide carrier (Mitochondrial uncoupling
DE protein 1).
GN SLC25A19 OR DNC OR MUP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX TISSUE=Liver;
RX PubMed=1126231;
RT Dolce V., Fiermonte G., Runswick M.J., Palmieri F., Walker J.E.;
RT "The human mitochondrial deoxynucleotide carrier and its role in the
RT toxicity of nucleoside antivirals.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:2284-2288(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Renard S., Mondesert G., Besnard F.;
RT "MUP 1, a mitochondrial uncoupling protein.";
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Responsible for the uptake of deoxynucleotides into the
CC matrix of the mitochondria. Transports all four deoxy NDPs, and,
CC less efficiently, the corresponding dNTPs. Does not transport
CC dNMPs, NMPs, deoxynucleosides, nucleosides, purines, or
CC pyrimidines. Supply deoxynucleotides to the mitochondrial matrix
CC for conversion to triphosphates and incorporation into
CC mitochondrial DNA.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined except for
CC placenta. Highest levels in colon, kidney, lung, testis, spleen,
CC and brain.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- DISEASE: Likely to be medically important by providing the means
CC of uptake into mitochondria of nucleoside analogs, leading to the
CC mitochondrial impairment that underlies the toxic side effects of
CC such drugs in the treatment of viral illnesses, including AIDS,
CC and in cancer therapy.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ251857; CAC27560.1; -.
```

```

DR EMBL; AJ301616; CAC37793.1; -.
DR EMBL; AF182404; AAG16903.1; -.
DR EMBL; BC001075; AAH01075.1; -.
DR EMBL; BC005120; AAH05120.1; -.
DR MIM; 606521; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
SQ SEQUENCE 320 AA; 35511 MW; 57CE0F01D538B1BE CRC64;

Query Match 12.9%; Score 90; DB 1; Length 320;
Best Local Similarity 24.6%; Pred. No. 0.031;
Matches 30; Conservative 25; Mismatches 33; Indels 34; Gaps 7;

Qy 7 DPAARYRNVLALWRIIRTEG---LWRPMRGLNVTATGAGPAHALYFACYEKLKLTSDV 63
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 54 DPSAKYHGILQASRGILQEEGPTAFWKGHVPAQILSIGYGAQVQLSF-----EMLTEL 106

Qy 64 IHPGG-----NSH-IANGAGCVATLLHDAAMP-----AEGND---SSTYHSV 103
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 107 VHRGVYDAREFSVHFVCGGLAACMATL-----TVHPVDVLTFRFAAQGEKPVYNTLRHAV 162

Qy 104 GS 105
Db 163 GT 164

RESULT 7
PHAB_CYACA
ID PHAB_CYACA STANDARD; PRT; 161 AA.
AC Q9TLS8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB.
OS Cyanidium caldarium.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RK-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.;
RT "The structure and gene repertoire of an ancient red algal plastid genome."
RT J. Mol. Evol. 51:382-390(2000).
CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022186; AAF12903.1; -.

```

```

DR HSSP; P00318; 1B33.
DR InterPro; IPR001659; Phycobilisome.
DR Pfam; PF00502; Phycobilisome; 1.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Methylation; Chloroplast.
FT MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE (BY
FT SIMILARITY).
SQ SEQUENCE 161 AA; 17471 MW; 3AFF46A2162AFDFA CRC64;

Query Match 12.9%; Score 89.5; DB 1; Length 161;
Best Local Similarity 28.3%; Pred. No. 0.015;
Matches 39; Conservative 18; Mismatches 52; Indels 29; Gaps 7;

Qy 7 DPAARY--RNVLALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKLT--SD 62
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 13 DVQGYLDNNSIEKLSRYFQT--GELRVRAAASISANAAG-----IIEAKVAKSLLYSD 64

Qy 63 VHPGCGNSHIANGAACGVATL---LHDAAMPNPAEGNDS-----STVHSV--S 105
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 65 ITRPGNMYTTRRYAACLRDLDYLYRYATYSMLAGDPSILDERVLNGLKETYNSLGVPIG 124
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Qy 106 CTCISLQFAEESTSVLVG 123
||:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 125 ATIQSIQAMKEVTASLVG 142

RESULT 8
PHAB_GALSU
ID PHAB_GALSU STANDARD; PRT; 161 AA.
AC P00319; P35910;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB.
OS Galdieria sulphuraria.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Galdieria.
OX NCBI_TaxID=130081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IIID-2;
RX MEDLINE=94269195; PubMed=8208852;
RA Liu B., Troxler R.F.;
RT "A Cyanidium caldarium allophycocyanin beta subunit gene."
RL Plant Physiol. 103:293-294(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=14-1-1 / ISOLATE 107.79/GOETTINGEN;
RA Kostorzewa M., Zetsche K.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE.
RX MEDLINE=83290919; PubMed=6885776;
RA Offner G.D., Troxler R.F.;
RT "Primary structure of allophycocyanin from the unicellular
RT rhodophyte, Cyanidium caldarium. The complete amino acid sequences of
RT the alpha and beta subunits."
RL J. Biol. Chem. 258:9931-9940(1983).
RN [4]
RP SEQUENCE OF 51-161 FROM N.A.
RC STRAIN=14-1-1 / ISOLATE 107.79/GOETTINGEN;
RX MEDLINE=94033298; PubMed=8219057;
RA Kostorzewa M., Zetsche K.;
RT "Organization of plastid-encoded ATPase genes and flanking regions
RT including homologues of infB and tsf in the thermophilic red alga
RT Galdieria sulphuraria."
RL Plant Mol. Biol. 23:67-76(1993).
CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM

```



```
[12]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=97061201; PubMed=8905231;
RA  Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA  Miyajima N., Hikosawa M., Suglura M., Sasamoto S., Kimura T.,
RA  Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA  Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA  Yamada M., Yasuda M., Tabata S.;
RT  "Sequence analysis of the genome of the unicellular cyanobacterium
RT  Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT  entire genome and assignment of potential protein-coding regions.";
RL  DNA Res. 3:109-136(1996).
[3]
RN  SEQUENCE OF 1-20.
RP  MEDLINE=97443974; PubMed=9298645;
RA  Sazuka T., Ohara O.;
RT  "Towards a proteome project of cyanobacterium Synechocystis sp.
RT  strain PCC6803: linking 130 protein spots with their respective
RT  genes.";
RL  Electrophoresis 18:1252-1258(1997).
CC  FROM THE PHCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS
CC  -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -|- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; M71135; AAA27277.1; -
DR  EMBL; D90910; BAA17875.1; -
DR  PIR; B44462; B44462.
DR  HSSP; P00318; 1B33.
DR  InterPro; IPR001659; Phycobilisome.
DR  Pfam; PF00502; Phycobilisome; 1.
KW  Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW  Methylation.
FT  MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT  BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE
FT  BINDING 81 81
SQ  SEQUENCE 161 AA; 17215 MW; 63F01E5903BA1B83 CRC64;
-----
Query Match 11.9%; Score 82.5; DB 1; Length 161;
Best Local Similarity 29.0%; Pred. No. 0.085;
Matches 27; Conservative 15; Mismatches 24; Indels 27; Gaps 4;
Qy 54 EKLKTL--SDVHPGGNSHTANGAGCV-----ATLLHDAAMNPAE 93
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54 EAVAKSLYSDVTRPGGNYTTRRYAACIRDLDDYLYRYATYAMLADGASILDRLVINGLK 113
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 94 GNDSTYHSVG---SCTCISLQFAEESTSVLGV 123
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 ----ETYNLSGLVPISTVQAIQAIKEVTASLVG 142
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
PHAB_SYNY4 STANDARD; PRT; 161 AA.
ID PHAB_FREDI
AC P16571;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB1.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
OX NCBI_TaxID=11197;
[1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=93222481; PubMed=8467079;
RA  Dinagno L.M., Haselkorn R.;
RT  "Isolation and characterization of the genes encoding allophycocyanin
RT  subunits and two linker proteins from Synechocystis 6714.";
RL  Plant Mol. Biol. 21:835-846(1993).
CC  -|- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC  FROM THE PHYCOCYANOPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC  -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -|- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; L02308; AAA69683.1; -
DR  HSSP; P00318; 1B33
DR  InterPro; IPR001659; Phycobilisome.
DR  Pfam; PF00502; Phycobilisome; 1.
KW  Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW  Methylation.
FT  MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT  BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE.
FT  BINDING 81 81
SQ  SEQUENCE 161 AA; 17242 MW; 63F0047008630A73 CRC64;
-----
Query Match 11.9%; Score 82.5; DB 1; Length 161;
Best Local Similarity 29.0%; Pred. No. 0.085;
Matches 27; Conservative 15; Mismatches 24; Indels 27; Gaps 4;
Qy 54 EKLKTL--SDVHPGGNSHTANGAGCV-----ATLLHDAAMNPAE 93
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54 EAVAKSLYSDVTRPGGNYTTRRYAACIRDLDDYLYRYATYAMLADGASILDRLVINGLK 113
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 94 GNDSTYHSVG---SCTCISLQFAEESTSVLGV 123
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 ----ETYNLSGLVPISTVQAIQAIKEVTASLVG 142
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
PHAB_FREDI STANDARD; PRT; 162 AA.
ID PHAB_FREDI
AC P16571;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB1.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
OX NCBI_TaxID=11197;
[1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=89053869; PubMed=2461358;
RA  Hounard J., Capuano V., Coursin T., Tandeau de Marsac N.;
RT  "Genes encoding core components of the phycobilisome in the
RT  cyanobacterium Calothrix sp. strain PCC 7601: occurrence of a
RT  multigene family.";
RL  J. Bacteriol. 170:5512-5521(1988).
CC  -|- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC  FROM THE PHYCOCYANOPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
CC  ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
CC  -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -|- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


SEQUENCE OF 55-82, AND METHYLATION.
 RA Ruemelli R., Suter F., Wirth M., Sidler W., Zuber H.;
 RT "Gamma-N-methylasparagine in phycobiliproteins from the cyanobacteria
 RL Mastigocladus laminosus and Calothrix.";
 RN FEBS Lett. 221:1-2(1987).
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RC STRAIN=PCC 7603;
 RX MEDLINE=99145555; PubMed=9990029;
 RA Reuter W., Wiegand G., Huber R., Than M.E.;
 RT "Structural analysis at 2.2 A of orthorhombic crystals presents the
 RT asymmetry of the allophycocyanin-linker complex, AP-1C7.8, from
 RT phycobilisomes of Mastigocladus laminosus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1363-1368(1999).
 CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
 CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM
 CC ABSORPTION AT APPROXIMATELY 650 NANOMETERS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
 DR PIR: S02501; AFMBW.
 DR PIR: S02612; S02612.
 DR PDB: 1B33; 23-FEB-99.
 DR InterPro: IPR001659; Phycobilisome.
 DR Pfam: PF00502; Phycobilisome; 1.
 KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
 KW Methylation; 3D-structure.
 FT MOD_RES 71 71 METHYLATION.
 FT BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE.
 FT CONFLICT 59 60 SL -> KT (IN REF. 1).
 FT CONFLICT 67 67 R -> L (IN REF. 1).
 FT CONFLICT 71 71 N -> D (IN REF. 1).
 SQ SEQUENCE 161 AA; 17374 MW; 18E7DD78D7178350 CRC64;

Query Match 11.4%; Score 79.5; DB 1; Length 161;
 Best Local Similarity 31.5%; Pred. No. 0.18;
 Matches 28; Conservative 12; Mismatches 30; Indels 19; Gaps 4;
 QY 54 EKLKKTLL--SDVTHPGNGSHIANGAAGCVATL---LHDAAMNPAEGNDS----- 97
 DB 54 EAVAKSLYSDFTRPGNMYTTRRYAACIRDLDYLYRYATYAMLAGDPSILDERVLNGLK 113
 QY 98 SYTHSVG---SCTCISLQFAESTSVLVG 123
 DB 114 ETYNSLGVPISTVQAIQAKVETASLVG 142

RESULT 15
 M2OM_HUMAN
 ID M2OM_HUMAN STANDARD; PRT; 313 AA.
 AC Q02978; 075537;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP).
 GN SLIC25A11 OR SLIC20A4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93091249; PubMed=1457818;
 RA Iacobazzi V., Palmieri F., Runswick M.J., Walker J.E.;
 RT "Sequences of the human and bovine genes for the mitochondrial 2-
 RT oxoglutarate carrier.";
 RL DNA Seq. 3:79-88(1992).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RA Yu W., Gibbs R.A.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE TRANSPORT OF 2-OXOGLUTARATE ACROSS THE

INNER MITOCHONDRIAL MEMBRANE IN AN ELECTRONEUTRAL EXCHANGE FOR
 CC MALATE OR OTHER DICARBOXYLIC ACIDS, AND PLAYS AN IMPORTANT ROLE
 CC IN SEVERAL METABOLIC PROCESSES, INCLUDING THE MALATE-ASPARTATE
 CC SHUTTLE, THE OXOGLUTARATE/ISOCITRATE SHUTTLE, IN GLUCONEOGENESIS
 CC FROM LACTATE, AND IN NITROGEN METABOLISM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X66114; CAA46905.1; -.
 DR EMBL: AF070548; AAC28637.1; -.
 DR PIR: S29598; S29598.
 DR MIM: 604165; -
 DR InterPro: IPR001993; Mitoch_carr.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT TRANSMEM 23 41 POTENTIAL.
 FT TRANSMEM 82 100 POTENTIAL.
 FT TRANSMEM 118 139 POTENTIAL.
 FT TRANSMEM 182 201 POTENTIAL.
 FT TRANSMEM 221 239 POTENTIAL.
 FT TRANSMEM 280 299 POTENTIAL.
 FT REPEAT 1 107 1.
 FT REPEAT 108 208 2.
 FT REPEAT 209 308 3.
 FT CONFLICT 11 11 I -> M (IN REF. 1).
 SQ SEQUENCE 313 AA; 33930 MW; A4831A2E1A9F175A CRC64;

Query Match 11.4%; Score 79.5; DB 1; Length 313;
 Best Local Similarity 31.1%; Pred. No. 0.39;
 Matches 28; Conservative 12; Mismatches 49; Indels 1; Gaps 1;
 QY 4 LQDPDAARYNVNLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKIKKTLSDV 63
 DB 152 LPADQRRGYKNVFNALIRITREEGVLTLMRGCIPTMARAVVNVVNAQAQLASYSQSKQFLDS 211
 QY 64 IHPGGNSHIANGAAGCVATLLHDAAMNPAE 93
 DB 212 GYFSDNI-LCHFCAISMISGLVTTAASMPVD 240

Search completed: August 27, 2002, 04:16:49
 Job time: 1241 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 04:19:24 ; Search time 130.92 seconds
(without alignments)
173.101 Million cell updates/sec

Title: US-09-870-113-12
Perfect score: 695
Sequence: 1 MQSLQPDPAARYRNVEALW.....QFAESTSVLCVNSVTLFVH 131

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriaph.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	505	72.7	364	4	Q96A46	Q96A46 homo sapien
2	351	50.5	347	4	Q9NYZ2	Q9NYZ2 homo sapien
3	348	50.1	338	11	Q920G8	Q920G8 mus musculus
4	280	40.3	155	4	Q969S1	Q969S1 homo sapien
5	276	39.7	182	11	Q91ZY0	Q91ZY0 mus musculus
6	216.5	31.2	312	5	Q23125	Q23125 caenorhabdi
7	178.5	25.7	379	5	Q9VAY3	Q9VAY3 drosophila
8	169	24.3	380	5	Q9NYH6	Q9NYH6 drosophila
9	168	24.2	303	5	Q94638	Q94638 onchocerca
10	166	23.9	301	5	Q94634	Q94634 onchocerca
11	143.5	20.6	781	10	Q9LMJ6	Q9LMJ6 arabidopsis
12	141.5	20.4	289	10	O82049	O82049 ribes nigru
13	139	20.0	331	10	O64731	O64731 arabidopsis
14	119	17.1	303	3	O14281	O14281 schizosacch
15	115	16.5	320	11	Q921P8	Q921P8 mus musculus
16	114	16.4	345	11	Q9CYJ1	Q9CYJ1 mus musculus

17	108	15.5	321	4	Q9BSK2	Q9BSK2 homo sapien
18	101.5	14.6	384	5	Q27244	Q27244 caenorhabdi
19	99.5	14.3	322	10	Q9ARL9	Q9ARL9 hordeum vul
20	99.5	14.3	413	5	Q9VWF9	Q9VWF9 drosophila
21	99	14.2	307	4	O14589	O14589 homo sapien
22	98	14.1	322	10	Q9AX03	Q9AX03 oryza sativ
23	96	13.8	338	4	Q9UHR1	Q9UHR1 homo sapien
24	95.5	13.7	310	4	Q9NVN5	Q9NVN5 homo sapien
25	95.5	13.7	311	4	Q96CQ1	Q96CQ1 homo sapien
26	95.5	13.7	311	11	Q922G0	Q922G0 mus musculu
27	95	13.7	296	10	Q9CA93	Q9CA93 arabidopsis
28	94.5	13.6	313	10	Q9SJY5	Q9SJY5 arabidopsis
29	94.5	13.6	313	10	Q94K32	Q94K32 arabidopsis
30	94.5	13.6	337	4	Q9P182	Q9P182 homo sapien
31	94.5	13.6	351	4	Q9Y379	Q9Y379 homo sapien
32	94.5	13.6	351	4	Q9UF66	Q9UF66 homo sapien
33	94.5	13.6	359	4	Q9BZJ4	Q9BZJ4 homo sapien
34	94	13.5	318	11	Q9DAM5	Q9DAM5 mus musculu
35	93.5	13.5	336	11	Q9D981	Q9D981 mus musculu
36	91	13.1	132	4	Q9H0G8	Q9H0G8 homo sapien
37	90	12.9	311	10	Q9C6K8	Q9C6K8 arabidopsis
38	90	12.9	320	4	Q9HC21	Q9HC21 homo sapien
39	90	12.9	363	10	Q9M2Z8	Q9M2Z8 arabidopsis
40	90	12.9	376	10	Q9FR17	Q9FR17 arabidopsis
41	89.5	12.9	359	11	Q9D8K8	Q9D8K8 mus musculu
42	89	12.8	310	5	Q18844	Q18844 caenorhabdi
43	89	12.8	348	10	Q9SH98	Q9SH98 arabidopsis
44	89	12.8	358	5	Q9VQ37	Q9VQ37 drosophila
45	89	12.8	365	5	Q95TJ5	Q95TJ5 drosophila

ALIGNMENTS

RESULT 1

ID	Q96A46	PRELIMINARY;	PRT;	364 AA.
AC	Q96A46;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DE	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	MITOCHONDRIAL RNA SPLICING PROTEIN 3/4 (PUTATIVE MITOCHONDRIAL SOLUTE CARRIER SPLICER VARIANT).			
GN	HMRS3/4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21195335; PubMed=11297739;			
RA	Li F., Nikali K., Grogan J., Leibiger I., Leibiger B., Schweyen R.,			
RA	Larsson C., Suomalainen A.;			
RT	*Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4.;			
RL	FEBS Lett. 494:79-84(2001).			
DR	EMBL; AJ303077; CAC27996.1;			
DR	EMBL; AF327402; AAK49519.1;			
SQ	SEQUENCE 364 AA; 39271 MW; 22049C4F8DD7A577 CRC64;			

Query Match 72.7%; Score 505; DB 4; Length 364;
Best Local Similarity 87.5%; Pred. No. 5.3e-45;
Matches 99; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

Qy	1	MQSLQPDPAARYRNVEALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 60	
Db	100	MQSLQPDPAARYRNVEALWRIITEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLL 159	
Qy	61	SDVTHPGGNSHIANGAGCVATLLHDAAMPAAE-----GNDSTYHSVSC 106	
Db	160	SDVTHPGGNSHIANGAGCVATLLHDAAMPAAEYVVKQMOMYNSPYHRVTDC 211	


```

QY 1 MOSLOPDPARYRNVLALRIITRTGLWPRMGLNVTATGAGPAHALYFACYEKLKTL 60
Db 73 MOSLPPDKARYSIYIGALKRIMHTTEGFWRLRGLNMMGAGPAHAMYFACYENMKRTL 132
QY 61 SDVHPGNGSHANG 75
Db 133 NDVFSHOGNSHLANG 147

RESULT 6
Q23125 PRELIMINARY; PRT; 312 AA.
AC Q23125;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE W02B12-9.01 (TREMBLrel. 19, Last annotation update)
DE W02B12.9.01 (TREMBLrel. 19, Last annotation update)
GN W02B12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP Swinburne J., Ainscough R.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP Swinburne J., Ainscough R.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z66521; CAAG1399.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 312 AA; 34093 MW; D65FC4DD2936F76A CRC64;

Query Match 31.2%; Score 216.5; DB 5; Length 312;
Best Local Similarity 50.0%; Pred. No. 1e-14;
Matches 47; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 MOSLOPDPARYRNVLALRIITRTGLWPRMGLNVTATGAGPAHALYFACYEKLKTL 60
Db 45 MOSLCPETKCPKTPVHSLMSIVKREGWRLRGVNAAGSMFPAHALYFTVYKMGYL 104
QY 61 SDVHPGNGSH-ANGAAGCVATLLHDAANPAE 93
Db 105 TG--NSAGHSNTLAYGASGVVATLIHDAINPAE 136

RESULT 7
Q9VAY3 PRELIMINARY; PRT; 379 AA.
AC Q9VAY3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG4963 PROTEIN (GH09840P).
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP Swinburne J., Ainscough R.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Worton J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
DR EMBL; AF003763; AAF56764.1; -.
DR EMBL; AY060268; AAL25307.1; -.
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Inner membrane; Mitochondrion; Transmembrane; Transport.
SQ SEQUENCE 379 AA; 41794 MW; F6A20BABA35F5C2E CRC64;

```

```

Query Match 25.7%; Score 178.5; DB 5; Length 379;
Best Local Similarity 39.3%; Pred. No. 1.3e-10;
Matches 44; Conservative 17; Mismatches 40; Indels 11; Gaps 3;

QY 1 MOSLOPDPARYRNVLALRIITRTGLWPRMGLNVTATGAGPAHALYFACYEKLKTL 60
Db 42 MOSL--SPTKNNMIVSTLTMTREGLLRPIRGASAVLGAAGPAHSLFAAEMTKELT 99
QY 61 SDVHPGNGSHANGAAGCVATLLHDAANPAE-----GNDSSTYHVSVC 106
Db 100 AKFTSVRNLYVISGA--VATLIHDAISSPTDVIKQRMQMYNSPTSVSC 148

```

```

RESULT 8
Q9NHV6
ID Q9NHV6 PRELIMINARY; PRT; 380 AA.
AC Q9NHV6
DT 01-FEB-1997 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-FEB-1997 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN.
GN CG4963.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Schmid K.J., Tautz D.;
RL "A screen for rapidly evolving genes from Drosophila.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217402; AAF73387.1; -
DR FlyBase; FBgn0039561; CG4963.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR002067; Mit_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 41844 MW; 4C9AA524B97F8C6C CRC64;

Query Match 24.3%; Score 169; DB 5; Length 380;
Best Local Similarity 38.1%; Pred. No. 1.3e-09;
Matches 43; Conservative 17; Mismatches 41; Indels 12; Gaps 4;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 MQSL--SPPTKNNMIVSTLRTMITREGLRLPIRGASAVVLGAGPTSLYFAAYETKLT 99
Qy 61 SDVIHPGNSHIANGAACGCVATLLHDAAMNPAE-----GNDSSP--YHSVCS 106
Db 100 AKFTSVRLNLYVIGA---VATLHDAISSPTDIKTAYADCTTRTPHPVVC 149

RESULT 9
Q94638
ID Q94638 PRELIMINARY; PRT; 303 AA.
AC Q94638
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RL "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Mitochondrial solute carriers.";
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45998; AAB19037.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 303 AA; 33861 MW; 224BFA547E5D617C CRC64;

Query Match 24.2%; Score 168; DB 5; Length 303;
Best Local Similarity 36.6%; Pred. No. 1.3e-09;

Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 LQSLCPCPETSCPTAMHSLMSVKEGLRLSLKGVNAVVLGTIPAHAFYTYVYENSKAYL 101
Qy 61 SDVIHPGNSHIANGAACGCVATLLHDAAMNPAE 93
Db 102 LN--NPRVSNVSVAISGALATVIHDAVMNPAE 132

RESULT 10
Q94634
ID Q94634 PRELIMINARY; PRT; 301 AA.
AC Q94634
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MITOCHONDRIAL SOLUTE CARRIER.
OS Onchocerca gibsoni.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6284;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96326580; PubMed=8703971;
RA Catmull J., Miller D.J.;
RL "cDNAs from Onchocerca sp. encoding members of the MRS3/MRS4 class of
RL Mitochondrial solute carriers.";
RL Biochim. Biophys. Acta 1282:179-181(1996).
DR EMBL; U45997; AAB19036.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
SQ SEQUENCE 301 AA; 34176 MW; C0BA8D819FB8EA79 CRC64;

Query Match 23.9%; Score 166; DB 5; Length 301;
Best Local Similarity 36.6%; Pred. No. 2.1e-09;
Matches 34; Conservative 20; Mismatches 37; Indels 2; Gaps 1;

Qy 1 MOSLOPDPAAARNVLEALWRIIRTEGLWRPMRGUNVTATGAGPAHALYFACYEKLKRTL 60
Db 42 LQSLCPCPETSCPTAMHSLMSVKEGLRLSLKGVNAVVLGTIPAHAFYTYVYENSKAYL 101
Qy 61 SDVIHPGNSHIANGAACGCVATLLHDAAMNPAE 93
Db 102 LN--NPRVSNVSVAISGALATVIHDAVMNPAE 132

RESULT 11
Q9LMJ6
ID Q9LMJ6 PRELIMINARY; PRT; 781 AA.
AC Q9LMJ6
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F10K1.26 PROTEIN.
GN F10K1.26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Liu S.X., Chan A., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Sakano H., Toriumi M., Chung M., Goldsmith A., Liu A., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RA "The sequence of BAC F10K1 from Arabidopsis thaliana chromosome 1.";

```

Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC067971; AAF82217.1; -
 DR InterPro: IPR001993; Mitoch_carrier.
 DR InterPro: IPR002067; Mit_carrier.
 DR Pfam: PF00153; mito_carr; 4.
 DR PRINTS: PR00926; MITOCHARRIER.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
 SQ SEQUENCE 781 AA; 87081 MW; 9FB579B9BD746D1E CRC64;

Query Match 20.6%; Score 143.5; DB 10; Length 781;
 Best Local Similarity 33.6%; Pred. No. 1.5e-06;
 Matches 49; Conservative 16; Mismatches 52; Indels 29; Gaps 6;
 QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKKTLL 60
 DB 64 MQALRCPCLKPVGIRAFRSIIQKEGPSALYRGWAGLGLGAGPAHAYFSFVSVSKYL 122
 QY 61 SDVIHPGG--NSHIANGAAGCVATLLHDAAMNP-----AEGNDSSTYHSVGSCT- 107
 DB 123 S-----AGDQNNVAHAMSGVFATISSDAVFTPMVMYKQLQMGEG----TYKGVWDCVK 173
 QY 108 -----CISLQFAESTSVLGNVST 127
 DB 174 RVLREGIGAFYASYRTTVMNAPFT 199

RESULT 12
 OS 82049 PRELIMINARY; PRT; 289 AA.
 AC 082049;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MITOCHONDRIAL CARRIER PROTEIN.
 GN PRIB7.
 OS Ribes nigrum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Saxifragales; Grossulariaceae; Ribes.
 ON NCBI_TaxID=78511;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
 RA Woodhead M.R.; Taylor M.A., Brennan R.M., McNicol R.J., Davies H.V.;
 RL Thesis (1995), University of Dundee, UK.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BEN ALDER; TISSUE=FRUIT;
 RA Woodhead M.R., Taylor M.A., Brennan R.M., McNicol R.J., Davies H.V.;
 RT "Cloning and characterisation of the cDNA clones of five genes that
 RT are differentially expressed during ripening in the fruit of
 RT blackcurrant (Ribes nigrum L.).";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL: AJ007580; CRA07568.1; -
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 KW Inner membrane; Mitochondrion; Transmembrane; Transport.
 SQ SEQUENCE 289 AA; 30383 MW; 5666EAB03DCC507C CRC64;

Query Match 20.4%; Score 141.5; DB 10; Length 289;
 Best Local Similarity 31.2%; Pred. No. 7.5e-07;
 Matches 40; Conservative 18; Mismatches 51; Indels 19; Gaps 3;
 QY 9 AARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKKTLLSDVIHPGG 68
 DB 32 SAQSAGLRQALGSLIKVEGPAGLYRGIGAGLGLGAGPAHAYFSVYEMKETFS---HGPD 88

QY 69 NSHIANGAAGCVATLLHDAAMNPAE-----GNDSTYHSVGSCTCISLQFAESTSVLV 122
 DB 89 SNSGAHAVSGVFATVSDAVITPMDVYKQLQLOSPKGVDCV-----RRVLV 138
 QY 123 GNSVTLFY 130
 DB 139 EEGIGAFY 146
 RESULT 13
 OS 064731 PRELIMINARY; PRT; 331 AA.
 AC 064731;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE MITOCHONDRIAL CARRIER PROTEIN.
 GN AT2G30160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004165; AAC16956.1; -
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mito_carr; 3
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
 SQ SEQUENCE 331 AA; 35961 MW; A0DE93084BBC8BC3 CRC64;

Query Match 20.0%; Score 139; DB 10; Length 331;
 Best Local Similarity 36.2%; Pred. No. 1.6e-06;
 Matches 42; Conservative 15; Mismatches 41; Indels 18; Gaps 5;
 QY 1 MQSLQDPDPAARYNVLEALWRIIRTEGLWPRMGLNVTATGAGPAHALYFACYEKLKKTLL 60
 DB 66 MQALRSCP-IKPIGIRQAFRSIIKTGDPGSAFYRGWAGLGLGAGPAHAYFSFVSVSKFL 124
 QY 61 SDVIHPGG--NSHIANGAAGCVATLLHDAAMNPAE-----GNDSTYHSVGSCT 106
 DB 125 S-----GGNPNSAAHAISGVFATISSDAVFTPMVMYKQLQIGN--GTYKGVWDC 173

RESULT 14
 OS 014281 PRELIMINARY; PRT; 303 AA.
 AC 014281;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PUTATIVE MITOCHONDRIAL CARRIER C8C9.12C.
 GN SPAC8C9.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:33:30 ; Search time 1936.36 Seconds
(without alignments)
11833.831 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atggagttggaggcgggg.....agtggaggctggcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_vt:*

14: gb_vt:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pi:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

RESULT 1

AF327402

LOCUS

DEFINITION

AF327402

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

ALIGNMENTS

AF327402

LOCUS

DEFINITION

AF327402

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

Key	Location/Qualifiers
FF	1..1448
FF	/db_xref="taxon:9606"
FF	/organism="Homo sapiens"
FF	30..1124
FF	/gene="HMR3/4"
FF	/product="mitochondrial RNA splicing protein 3/4"
FF	/function="putative mitochondrial solute carrier"
FF	/protein_id="CA27996.1"
FF	/translation="MELRGAGGVAGGPAAGPGRSGPESALLDGLQLRGVGRGAGGGE
FF	ACAGRPVVRQDPDSEALPAGATVTTHMVAGAVAGILEHCVMYRDCVKTMRQSLQ
FF	PPDARYRNVEALWRIIRTEGLWRPMRLNVTATGAGPAHALYFACYEKLKLTLSDI
FF	HPGNSHTAANGAAGCATLLHDAMNPAEVVQRQMYNSPYHRVTDVCRVAVQNEGAP
FF	AFYRSTLTQTMNVPFOALHPMTVEFLQEHFNQPRYRNPSPSHVLSGACAGAAAAATTP
FF	LDVCKTLNLTQSLALNSHITGHITGMASAFRTVYQGVGVTAIFYRQVQARVIYQIPSTA
FF	IANSVYEFKYLITKRQEWRAK"
XX	
QQ	Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;
QQ	Query Match 100.0%; Score 1095; DB 17; Length 1448;
QQ	Best Local Similarity 100.0%; Pred. No. 6.6e-186;
QQ	Matches 1095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 atgagttgagggcggggtgctggcggtgctggcgggggcgggcgagggcgccggg 60
QY	
QY	30 ATGAGTTGAGGGCGGGGTGTCGGGTGTGGCGGGGGCGCGCGAGGGCGCCGGG 89
QY	
QY	61 cggagcccgggagtcggcgctgctgtagcggtgagctgcagcgggcgtagggcgggg 120
QY	
QY	90 CGAGCCCCGGGAGTGGCGGTGTGTGACGGGTGGCTGCACGGGGCGTGGCGCGGGG 149
QY	
QY	121 gccggcggggagggcgggcgccgtgcagggccccgggtacgacaagatccggactccggc 180
QY	
QY	150 GCCGGCGGGGAGGCGCGGGCGCTGCAGGCCCGCCGGTACGACAGATCCGACTCCGGC 209
QY	
QY	181 ccggactcagagcgctgcggcgctggagccatgtccaccacacatggtggcagcgccc 240
QY	
QY	210 CCGGACTACAGGCGCTGCGCGCTGGAGGCCACTGTCAACCACACACATGTTGGCAGCGCC 269
QY	
QY	241 gtgcagggatcctggagcactcgctgtagtaccacctgactcgctcaagaccggatg 300
QY	
QY	270 GTGGCAGGATCTGTGAGGACTGCGTGATGTACCCCATCGACTGGCTCAAGACCCGATG 329
QY	
QY	301 cagagtctacgctgaccagctgcgcgctatcgcaatgtgttgaggccctctggagg 360
QY	
QY	330 CAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTTGGAGCCCTCTGGAGG 389
QY	
QY	361 attataagacaggagcgctatggagcccatgagggggctgaaagctcacagcaaacg 420
QY	
QY	390 ATTATTAAAGAACGGAGGGCTATGTGAAGGCCATGTAGGGGGCTGAACGTCAACAGCAAGCC 449
QY	
QY	421 gcagggcctgcccaagccctttatttgcctgctacgaaaaagttaaaaaacacattgagt 480
QY	
QY	450 GCAGGGCCTGCCACGCGCTTTATTTTGCTGCTACGAAAAGTTAAAAAGACATTGAGT 509
QY	
QY	481 gatgtaatccaccctgggggcaatgacccatattgccaatggtgcggcgcggtgtgtggca 540
QY	
QY	510 GATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGTTGGTGGCGCGGGTGTGTGCA 569
QY	
QY	541 acattactcatgatgcagccaatgaacctgcgggaagtggtaagcagagatgcagatg 600
QY	
QY	570 ACATTACTTCATGTACGCCATGAACCTCGGAAGTGGTCAAGCAGAGATGCAGATG 629
QY	
QY	601 tacaactcaccataccaccgggtgcagactgttacggcgagtggtgcgaatgaagg 660
QY	
QY	630 TACAACCTACCATACCCGGGTGACAGACTGTGTACGGGCAGTGTGCAAAATGAAGGG 689
QY	
QY	661 gccggggccttttacgcagctacaccaccagctgaccatgaagcttccttccaaagcc 720
QY	
QY	690 GCCGGGCGCTTTTACCGAGCTACACACCCAGCTGACCATGAACGTTCTTCCAAAGCC 749
QY	


```
|||||
Db 309 CAGAGTCTACAGCCTGACCCAGCTGCCCGCTATGCAATGTTGGAGGCCCTCTGGAGG 368
Qy 361 attataaagcaggagcctatgagggcccatgaggggctgaacgtcacagcaacaggc 420
Db 369 ATTATAAAGCGGAGGGCCTATGGAGGCCCATGAGGGGGCTGAACGTCAACAGACAGC 428
Qy 421 gcaggcctgccacgcctttatttgcctgctacgaaaagttaaaaaagacattgagt 480
Db 429 GCAGGGCCTGCCACGCCCTTATTTTGCCTGCTACGAAAAGTTAAAAAGACATTGAGT 488
Qy 481 gatgtaatccacctgggggcaatagccatatgccaat----- 519
Db 489 GATGTAATCCACCTGGGGCAATAGCCATATGTCGAATGGTATTGAGCCTTCCTGTGCT 548
Qy 520 ----- 519
Db 549 GGTCCCCCACTTTCCCAACTTTTGGGCTTTGTGCTGTGCTGCTCAGTCTTCCAGTCTCAGC 608
Qy 520 ----- 519
Db 609 ATGGTTGGAGCTGAAGCTTTGGGCTGGGATAGGCCACTGGTGCGAGGTGATGCTCGC 668
Qy 520 ----- 519
Db 669 ATCCCTAGCAGCAGTATCTGCTTCCTGCTGGGGTGTGAGCTGCATTTATTCTCAGA 728
Qy 520 ----- 519
Db 729 ATGATCTTTATGATAAGACTGAGCTGGCCTTCCTATCATCGGATGTGGAATACATTAGT 788
Qy 520 ----- 519
Db 789 GACCTTACAAAGTTGGTGGGAACAGATACTTTACCTTCTTAAACAGAGTATTAGGAGCAG 848
Qy 520 ----- 519
Db 849 TGGGTCCCCATCTTTTGGACTAGCTCTTAACGTTACTTTTCCCGCTGTAGTGTAGCACA 908
Qy 520 ----- 519
Db 909 GCCACTCCCTTCACTGGGGACCTCAGTGAGTTGGTCAGCTCTCTTGGCCTTACATGTG 968
Qy 520 ----- 519
Db 969 GCAGTTGTTTTCTTGTGTCAGGTGCGGCGGGTGTGTGGCAACATTACTTTCATGATGCA 1028
Qy 559 gccatgaacctgcggaagtggtcaagcagagagatgcagatgtacaactcaccataccac 618
Db 1029 GCCATGAACCCCTGCGGAAGTGTCAAGCAGAGAGGTGCAGATGTACAACCTCACCATACCAC 1088
Qy 619 cgggtgacagactgttacggcagtggtgcaaatgaaggcgccgggcccctttaccgc 678
Db 1089 CGGGTGACAGACTGTGTACGGCAGTGTGGCAAAATGAAGGGCGCGGGCCCTTTACCGC 1148
Qy 679 agtacacacccagctgaccatgaacgttccttcccaagccattcaactcatgacctat 738
Db 1149 AGCTACACCCAGCTGACCTACCATGAACGTTCTCTTCCAAAGCCATTACATTCATGACCTAT 1208
Qy 739 gaattcctcagagacactttaacccacagacaggtacaaccaagctccacgctccctc 798
Db 1209 GAATTCCTGCAGAGACATTTTAACCCCCAGAGCGGTACAACCCAGCTTCCACAGCTCCCTC 1268
Qy 799 tctggagcttgcgagagagctgtagtgcgcgagccacacacccacctggaacttttgcataa 858
Db 1269 TCTGGAGCTTGGCAGGAGCTGTAGTGCAGCCAGCCACAAACCCACTGAGGCTTTGCAAA 1328
Qy 859 acaactgctcaacacccagagctccttggctttgaactcacacattacaggaacatatca 918
Db 1329 ACACCTGCTCAACACCCAGAGAGCTTGGCTTTGAACCTCACATTAACAGGACATATCACA 1388
Qy 919 ggcattgctagtccttcagagcggtatatcaagtagtggtgggtaccgcctatttccga 978
|||||
```

```
Db 1389 GGCATGGCTAGTGCCTTCAGGACGGTATATCAAGTAGTGGGTGACCGCCTATTTCGGA 1448
Qy 979 ggggtgcaggccagagtaattaccagatccctccacagccatgcgatggtctgtgtat 1038
Db 1449 GGGGTGCAGGCGCAGAGTAATTTACCAGATCCCTCCACAGCCATCGCATGCTGTGTAT 1508
Qy 1039 gagttcttcaataacctaatcactaaaggaaggaagagtgagggtcggaagtga 1095
Db 1509 GAGTTCTTCAAATACCTAATCAGTAAAAAGCAAGAGAGTGGAGGGCTGGCAAGTGA 1565

RESULT 6
AF267854 Homo sapiens NP016 mRNA 1244 bp linear PRI 02-JAN-2001
LOCUS AF267854
DEFINITION Homo sapiens NP016 mRNA, complete cds.
ACCESSION AF267854
VERSION AF267854.1 GI:12006034
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1244)
AUTHORS Xu,X., Yang,X., Gao,G., Xiao,H., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China
LOCATION/Qualifiers
FEATURES
Source 1. 1244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="normal pituitary"
381..911
/codon_start=1
/product="NP016"
/protein_id="AAG44723.1"
/db_xref="GI:12006035"
/translation="MALLVKORMQNTSPYHRVTDVRAVWQNEGAGAFYRSYTTQL
TMNYPFOAIHFTYFLOEHFNPPORRYNPSHVSACAGAVAAAATPLDVCFTLLN
TQESIALNSHITGHITMASAFRTYQVGGVTAYFRGVQARVIYQIPSTAIANSVYEF
FKYLITKROEWRAGK"
BASE COUNT 308 a 331 c 290 g 315 t
ORIGIN
Query Match 47.5%; Score 520; DB 9; Length 1244;
Best Local Similarity 100.0%; Pred. No. 2.8e-83;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 576 agtgtcaagcagagagatgcagatgtacaactcaccataccacccgggtgacagactgtgt 635
Db 392 AGTGTCAAGCAGAGAGATGCAGATGTACAACCTACCACCGGGTGACAGACTGTGT 451
Qy 636 acggcagtgtagcaaatgaaggcgccggttttaccgcagctacacacccagct 695
Db 452 ACGGCAGTGTGGCAAAATGAAGGGCGCGGGCCCTTTACCGCAGCTACACCCAGCT 511
Qy 696 gaccatgaactgttcctttccaaagccattcaactcatgacctatgaattcctgcaggagca 755
Db 512 GACCATGAACGTTCTCTTCCAAAGCCATTCACTTCATGACCTATGAATTCCTCGAGGACA 571
Qy 756 ctttaaacccccagagacggtacaaccaagctccacgctcctctctggagcttgcgcagg 815
Db 572 CTTTAACCCCCAGAGCGGTACAACCCAAAGCTCCCAAGCTCTCTCTGAGAGCTTGGCAGG 631
Qy 816 agctgtagtgcgcgagcagccacacccacctggacctgttgcataaacactgctcaacacca 875
Db 632 AGCTGTAGTGTGGCAGCGCACACCCCACTGGAGCGTTTGCAAAACACTGCTCAACACCCA 691
Qy 876 ggagtccttggcttggaaactcacacattacaggaacatatcacaggcatggcttagtcctt 935
Db 692 GGAGTCTCTGGCTTTTGAACCTCACATTAACAGGACATATCACAGGATGGCTAGTGCCTT 751
```

Qy 936 caggacgttatcaagttagtggtgggtgaccgcctatttcgaggggtgcagggccagagt 995
 |||||||
 Db 752 CAGGACGGTATATCAAGTAGTGGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGT 811
 |||||||
 Qy 996 aattaccagatccctccacagccatgcgatgctgtgtatgatgttcttcaaatacct 1055
 |||||||
 Db 812 AATTACAGATCCCTCCACAGCCATCGCATGCTGTGTATGAGTTCTTCAAAATACCT 871
 |||||||
 Qy 1056 aatcactaaaggcaagaagtgaggcgtggaagtga 1095
 |||||||
 Db 872 AATCACTAAAAGGCAAGAAGTGGAGGGCTGGCAAGTGA 911
 |||||||

RESULT 7
 AL353719/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete sequence.
 ACCESSION AL353719 AC007643
 VERSION
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 123160)
 Ramsay, H.
 Direct Submission
 Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonequery@sanger.ac.uk

On Sep 26, 2001 this sequence version replaced gi:14280413.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr10
 RP11-85A1 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-85A1 it may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true left end of clone RP11-85A1 is at 1 in this sequence. The true left end of clone RP11-483F11 is at 123061 in this sequence. The true right end of clone RP11-129J12 is at 51589 in this sequence.

FEATURES

Location/Qualifiers
 1. .123160
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-85A1"
 /clone_lib="RPCI-11.1"

misc_feature complement(7066..7118)
 /note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
 52162
 unsure
 BASE COUNT 33221 a 28084 c 27196 g 34659 t
 ORIGIN

Query Match 47.5%; Score 520; DB 9; Length 123160;
 Best Local Similarity 100.0%; Pred. No. 1.6e-83;
 Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 576 agtgtcaagcagagatgcagatgtacaactcaccatcacccgggtgcagactgtgt 635
 |||||||
 Db 82990 AGTGGTCAAGCAGAGGATGCAGATGTACACTACCATACCCGGGTGCAGACTGTGT 82931
 |||||||
 Qy 636 acggggcagtggtggcaaatgaaggggggcgtttttaccgcagctcacacccccagct 695
 |||||||
 Db 82930 ACGGCGAGTGTGGCAAAATGAAGGGCGCGGCGCTTTTACCGCAGCTACACCCAGCT 82871
 |||||||
 Qy 596 gaccatgaacgttctcttccaaagccattcattcatgacctatgaattctctgcaggagca 755
 |||||||
 Db 82870 GACCATGAACGTTCTTTCCAAAGCCATTACCTTTCATGACCTATGAATTCCTGCAGGAGCA 82811
 |||||||
 Qy 756 ctttaacccccagagcaggtacaaacccaagctccacagctctctctgagcttgcgcaag 815
 |||||||
 Db 82810 CTTTAAACCCCGAGAGCGGTACAAACCCCAAGCTCCACGCTCTCTCTGGAGCTTGGCGAGG 82751
 |||||||
 Qy 816 agtgtgtagtcgcccagccacacacccactggacgttttgcaaaacactgctcaacaccca 875
 |||||||
 Db 82750 AGCTGTAGTGCCTGAGCAGCACACCCACCTGCAGGTTTGCAAAACACTGCTCAACACCCA 82691
 |||||||
 Qy 876 ggagtccttggttgactcaacattacagagcatatcacagggcgtgtagtgctt 935
 |||||||
 Db 82690 GGAGTCTCTGGCTTTGAACTCACACATACAGGACATATCACAGGCATATTCAGGCTTGGCTT 82631
 |||||||
 Qy 936 caggacgttatcaagttagtggtgaccgcctatttcgaggggtgcagggccagagt 995
 |||||||
 Db 82630 CAGGACGGTATATCAAGTAGTGGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGT 82571
 |||||||
 Qy 996 aattaccagatccctccacagccatgcgatgctgtgtatgatgttcttcaaatacct 1055
 |||||||
 Db 82570 AATTACAGATCCCTCCACAGCCATCGCATGCTGTGTATGAGTTCTTCAAAATACCT 82511
 |||||||
 Qy 1056 aatcactaaaggcaagaagtgaggcgtggaagtga 1095
 |||||||
 Db 82510 AATCACTAAAAGGCAAGAAGTGGAGGGCTGGCAAGTGA 82471
 |||||||

RESULT 8

AK056782
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ32220 fis, clone PLACE6003850, moderately similar to Homo sapiens mitochondrial solute carrier mRNA.
 ACCESSION AK056782
 VERSION
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE6 clone:PLACE6003850.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,

Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1588)
 Isogai, T., Otsuki, T. and Sugiyama, T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

FEATURES
 source
 1. .1588
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE6003850"
 /tissue_type="placenta"
 /clone_lib="PLACE6"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 385 a 405 c 375 g 423 t
 ORIGIN

Query Match 47.3%; Score 518.4; DB 9; Length 1588;
 Best Local Similarity 99.8%; Pred. NO. 5.2e-83;
 Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 576 agtgcagcagagatgcagatgatacaactcaccataccaccgggtgacagactgtgt 635
 Db 745 AGTGGTCAGCAGAGATGCAGATGTACACTCACCATTACCACCGGGTGCAGGCTGTGT 804
 QY 636 acgggcagtggtgcaaaatgaaggggccggccctttaccgcagctacacacccagct 695
 Db 805 ACGGGCAGGTGTGGCAAAATGAAGGGGCGGGGCCCTTTTACCAGCTACACCACCCAGCT 864
 QY 696 gaccatgaacgtctcttcaagcattcaactcctatgaactatgattcctcaggaagca 755
 Db 865 GACCATGAACGTTCTCTTCCAGGCCATTACATTCATGACCTATGAATTCCTCGCAGGACA 924
 QY 756 ctttaacccccagagcaggtatacaaccccaagctcccacgctcctctctggagcttgcgcagg 815
 Db 925 CTTTAACCCAGAGACGGTACAAACCAAGCTCCACGCTCTCTGGAGCTTGGCAGG 984
 QY 816 agctgtagctgcgcagccacacccccactgagcgttttgcaaaacactgctcaaacacca 875
 Db 985 AGCTGTAGTGGCGGACGACACACCCCACTGGACGCTTTGCAAAACACTGCTCAACACCCA 1044
 QY 876 gtagctctgcttgaactcacattacagacacatatcacagcagcatggtcagcttgcctt 935
 Db 1045 GGAGTCCTTGGCTTTGAACTCACATTCACAGACATATCACAGGCATGCGCTAGTGCCTT 1104
 QY 936 caggacggtatatacagtagtggtgggtgaccgcctatttccgaggggtgcagccagagt 995
 Db 1105 CAGGACGGTATATCAAGTAGTGGGTGACCGCCCTATTTCCGAGGGGTGCAGCCAGAGT 1164
 QY 996 aattaccagatccctccacagccatgcagctggtctgtgtatgagtttctcaataacct 1055
 Db 1165 AATTTACCAGATCCCTCCACAGCCATCGATGGTCTGTGTATGAGTTCTTCAATAACCT 1224
 QY 1056 aatcaactaaagcgaagagagtgaggctgcgaagtga 1095
 Db 1225 AATCACTAAAGGCAAGAGAGTGGAGGCTGGCAAGTGA 1264

RESULT 9
 BC023172
 LOCUS 1072 bp mRNA linear ROD 07-FEB-2002

DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT

Mus musculus, Similar to putative mitochondrial solute carrier,
 clone MGC:37028 IMAGE:4949779, mRNA, complete cds.
 BC023172
 BC023172.1 GI:18606247
 MGC.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1072)
 Strausberg, R.
 Direct Submission
 Submitted (04-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 60 Row: h Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source

1. .1072
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:37028 IMAGE:4949779"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma 5 month old virgin mouse."
 /clone_lib="NCI-CoAP_Mam6"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 188. .721
 /codon_start=1
 /product="Similar to putative mitochondrial solute
 carrier"
 /protein_id="AAH23172.1"
 /db_xref="GI:18606248"
 /translation="MNPAAVYVKQRMQNSPYHRVTCVRAVWQNEGAFYRSYTTQ
 LTMVFFQAIHFMTYEFLOEHFNQRYNPSHVLGACAGADAATPTLDVCKTL
 NTOESLANSNITGHITGMASAFRTVYQGVGTAFVGVQARVIYQIPSTAIAWSVY
 FFKYLITKQEEWRAGK"
 BASE COUNT 268 a 308 c 263 g 233 t

CDS

BASE COUNT
 ORIGIN

Query Match 46.0%; Score 504; DB 10; Length 1072;
 Best Local Similarity 92.2%; Pred. No. 2e-80; Mismatches 45; Indels 0; Gaps 0;
 Matches 531; Conservative 0;

QY 520 ggtcgccgggtgtgtggcaacattactctcatgagccatgaacctcgcggaagt 579
 Db 146 GGTGACGGGATGTGTGGCGACATTACTTTCATGATGAGCCATGAATCCACGCGAAGTC 205
 QY 580 gtcaagcagagatgcagatgtacaactcacaccaccggggtgacagactgtgtacgg 639
 Db 206 GTCAAGCAGAGATGCAGATGTACAACCTCGCGGTACACCGCGTGCAGACTGTGTTCGG 265

Qy	640	gcagtgtgcaaaatgaagggccggggccttttacgcgagctacacacccagctgacc	699
Db	266	GCATGTGGCAAAATGAAGGGCCGGGGCTTTTACCGCAGCTACAGACCCAGCTGACC	325
Qy	700	atgaacgtctcttcccaagccattcactcatgacctatgaattcctgcaggagcaacttt	759
Db	326	ATGAATGTCCCTTCCAAGGCATTCATTCTATGACCTATGAGTTCCTGCAAGACGACTTT	385
Qy	760	aacccccagagcgggtacaaacccaagctcccaactctctctctgtgagtttgcgcaggagct	819
Db	386	AACCCCCAGAGAGCGGTACAACCCAGCTGCCATGTGCTGTGGGGCTGCGCAGGAGCA	445
Qy	820	gtagctccgcagccacaaaccccaactgagcgtttgcaaaacactgctcaacacccagag	879
Db	446	GTAGCTGACGCGGCCACACCCCACTGGAGCTTTGTAAACACTGCTCAACACCCACGAA	505
Qy	880	tccttggctttgaactcacacattacaggacatatcacaggcatcgtgtagtccttcagg	939
Db	506	TCCCTGGCTTTGAACCTCAAACATTTACTGGACACATCACAGGCATGGCTAGTGCTTCAGG	565
Qy	940	acggtatatcaagttagtggggtagccgcctatttcgaggggggtgcaggccagagttaatt	999
Db	566	ACGCTCTATCAAGTAGGGGGGTGACTGCTTACTTCCGAGGGGTGCAGGCCAGAGTCATT	625
Qy	1000	taccagatccctccacagccatcgcatcggtcgtgtgatgagttctcaataacctaatc	1059
Db	626	TACCAGATCCCTCCACAGCCATCGCATGGTCTGTGTATGTAATTTTTCAAATACCTATC	685
Qy	1060	actaaaggcgaagagtgtagggcctgcgaagtga	1095
Db	686	ACAAAGCGGCAAGAGGAATGGAGGCGAGGCAAGTGA	721

RESULT	10
AC096351/c	
LOCUS	221062 bp DNA linear HTG 20-DEC-2001
DEFINITION	Rattus norvegicus chromosome Rfl clone CH230-24W6, WORKING DRAFT SEQUENCE, 33 unordered pieces.
ACCESSION	AC096351
VERSION	AC096351.2 GI:17944054
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 221062)
REFERENCE	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carlon,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dunn-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsy,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegue,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Nowshat,N.,

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT

Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
 Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
 Quiles,M., Ren,Y., River,S., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
 Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Varra,V., Villalon,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 221062)
 Worley,K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:15627972.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

 Center project name: GE2D
 Center clone name: CH230-24M6
 ----- Summary Statistics

 Assembly program: Phrap; version 0.990329First call to
 findPhrapList
 Consensus quality: 200164 bases at least Q40
 Consensus quality: 205303 bases at least Q30
 Consensus quality: 209366 bases at least Q20
 Estimated insert size: 209880; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1
 * 32851: contig of 32851 bp in length
 * 32852: gap of unknown length
 * 32951: contig of 16033 bp in length
 * 32952: gap of unknown length
 * 48985: contig of 14731 bp in length
 * 49085: gap of unknown length
 * 63816: contig of 12090 bp in length
 * 63916: gap of unknown length
 * 76005: contig of 12736 bp in length
 * 76006: gap of unknown length
 * 76105: contig of 11650 bp in length
 * 88841: contig of 11011 bp in length
 * 88842: gap of unknown length
 * 88942: contig of 10862 bp in length
 * 100591: contig of 6947 bp in length
 * 100592: gap of unknown length
 * 100692: contig of 7436 bp in length
 * 111703: gap of unknown length
 * 111803: contig of 10862 bp in length
 * 122665: gap of unknown length
 * 122765: contig of 6947 bp in length
 * 129711: gap of unknown length
 * 129712: contig of 7436 bp in length
 * 129812: gap of unknown length
 * 132448: gap of unknown length
 * 137348: contig of 7329 bp in length
 * 144677: gap of unknown length
 * 152124: contig of 7348 bp in length
 * 144777: gap of unknown length


```
Db 164 ATGATGGGACTGCCGGATGGCGGTGTCGGCAGCAAGGACGCGGTGCGGAGGACTAC 223
Qy 190 gagcgctgcgggtgagccactgtcaccaacacacatggtgagcagcgccgtgagcggg 249
Db 224 GAGAACTACCGACCGCGCTCGGTGTCACACACACATGACGCGGAGCGATGCCGGG 283
Qy 250 atcctgagcactcgtgatgtacccatgcactgcgtcaagaccgcgagtcagagacta 309
Db 284 ATTCTGAGCACTCCATCAATGATCCCGGTGACATCCGTGAAGACACGATGACAGTTG 343
Qy 310 cagcctgaccagctgcccctgcgaatggtgtgagggccctctgagagctataaga 369
Db 344 AATCCAGATCCAAAGCCCGGTATACAGCATCTATGCGCGCCCTCAAGAGGATCATGCAC 403
Qy 370 acgagggccctatgagggcccatgaggggctgaagctcacagacaacagcgagggcct 429
Db 404 ACTAAGGCTTCTGGAGGCGCCCTCGGGGCGCTGAACGTGATGATGGGTGCGAGGCCC 463
Qy 430 gccacgcccctttatttgcctgtacgaaagttaaataaacacattgagtgatgaatc 489
Db 464 GCGACGCCATGATTTTGGCTGTATGAACACATGAAGGACCTTAATGACGTTTTC 523
Qy 490 caccctggggcaatagccattatgccaatggtgcggccgggtgtgtgcaacattactt 549
Db 524 AGCCACCAAGGAACAGCCCATCTAGCTAATGGGTAGCTGGGAGTATGGCCACCCTACTC 583
Qy 550 catgatcagccatgaacccctgcgaagtgtcaagcagagagatgcagatgataactca 609
Db 584 CACGATCAGTAATGAATCAGCAGAGAGTGTGAACACGCGCTTACAGATGTACAACTCC 643
Qy 610 ccatcaccacgggtgcagagactgtgcagggcagtggtggcaaatgaagggccggggcc 669
Db 644 CAGCACGATCAGCCTTCAGTTGTATCCGGACAGTGTGGCGGACCGAGGGTTGGGGCC 703
Qy 670 tttaaccgagctacacccacccagctgaccatgaaccttcccttcccaagccattcattc 729
Db 704 TTCTACAGGAGTTACACACACAGCTGACCATGAATATCCCTTCCAGTCAATTCACCTC 763
Qy 730 atacattataattctgcagagcactttaacccccagagaggtacaacccaagctcc 789
Db 764 ATACCTATGAGTTCTGCAAGGAGCAAGTCAACCCCTCGCCGGGACTACACCCACAGCTC 823
Qy 790 cagctctctctgaggttcgcagaggagctgagctgcgcagccacacaccccaactggac 849
Db 824 CACATCATCTCAGGAGGCGCTGGCGGGGCACTGGCCGCACTGCCACCCCGCTGGAC 883
Qy 850 gtttgaaaaactgctcaacacccagaggtccttggttgaaact---cacacattaca 906
Db 884 GTCTGCAAAACCTCTCAACACGCAAGAGAACATGCTCTCTCCCTGGCCACGTCAGC 943
Qy 907 ggacatacacagcagtggttagtccttcagcaggtatataaagttagtggtggtgacc 966
Db 944 GGCCGGCTGTGGGCATGGCCATGCCCTTCGGACGGGTATACAGCTCAACGGCCTTGCC 1003
Qy 967 gctatttcgaggggtgcagggccagagtaattaccagatccccctccacagccatcgca 1026
Db 1004 GGCTATTTCAAGGATCCAGGCTCGAGTCACTTACCAGATGCTTCCACCGCATCTCC 1063
Qy 1027 tggctgtgatgagttcttcaataactaatcaataaaaggcaagaag 1077
Db 1064 TGGTCCGCTTTATGAGTCTTCAAGTACATCTTACAAAGAGGCGAGCTGGAG 1114
```

RESULT 12

```
AF223466 AF223466 1429 bp mRNA linear PRI 17-APR-2000
LOCUS Homo sapiens HT015 protein (HT015) mRNA, complete cds.
DEFINITION
ACCESSION AF223466
VERSION AF223466.1 GI:7578782
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1429)
Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z., and Han,Z.
A novel gene expressed in human hypothalamus
Unpublished
2 (bases 1 to 1429)
Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z., and Han,Z.
Direct Submission
Submitted (12-JAN-2000) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, China
FEATURES
source
1..1429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
1..1429
/gene="HT015"
45..1088
/gene="HT015"
/codon_start=1
/product="HT015 protein"
/protein_id="AAF64141.1"
/db_xref="GI:7578783"
/translaton="MELRSQSVQAVARMGDSDRGSGKDATGSEYENLPTGAS
VSTHMTAGAMAGILEHSVMYPVDSVKRMOSLPDKAOTSYIGALKKIMRTEGFWR
PLRGVNVIMINGAGPAHAMYFACYENMKRLINDVPHQGNHSLANGLAGSNATLLHDV
MNPAYVKORLMYNPOHRSALSCIRTVWTEGIGAFYRITITQLTWNIPFQSIHFIT
YAEFQVNPHTYNPOSHIISGLAGALAAATPLDVKTLNQLTQENVALSNIS
GRLSGMANAFRTYQLNGLPATSKASRRVSSSTRCPPPPFLGLSSSSSTFSPSASWKI
ELHTKGRDHRIFS"
BASE COUNT 326 a 426 c 402 g 275 t
ORIGIN
Query Match 39.9%; Score 436.4; DB 9; Length 1429;
Best Local Similarity 66.4%; Pred. No. 2.3e-68;
Matches 673; Conservative 0; Mismatches 336; Indels 5; Gaps 3;
Qy 61 cggagcccccgggagtcggtgcgtgcgtgcagcgggtgcgtgcagcgggtgcgtgcggggg 120
Db 23 CTGCGCCCCCGGAGCTGGCGGATGGAGCTGCGCAGCGGGAGCGTGGCGCAGCCAGCGCGT 82
Qy 121 gccggcgccgggagggcggtgcgtgcgtgcagcccccggtacgacaagatccggactccgc 180
Db 83 GCGCGGAGGATGGATGGGACACGCCAGATGGCGGGCGGGCAAGGACGCCACCGGGTC 142
Qy 181 -ccggactacgagcgcgtgcgcgtgcgagccactgtcaccacgcacatggtgagcagcgc 239
Db 143 GGAGGACTACGAGAACCTCGCGACTAGCGCTCGGTGTCACACCCACATGACAGCAGGAGC 202
Qy 240 cgtgcagggatcctgagcactcgtgatgtaccccatcagactgcgtcgaagaccggat 299
Db 203 GATGGCCGGGATCCTGGAGCAGCTCGGTCAATACCCGGTGGACTCGGTGAAGACCAAT 262
Qy 300 gcagagtctacagctgcacccagctgcctcgttcgcaatggtgttgagggcctcttgag 359
Db 263 GCAGAGTTTGAGTCCAGATCCCAAAGCCAGTACACAAATATCTACGAGGCCCTCAAGAA 322
Qy 360 gattataaagcagggggccctatggagggccatagggggctggaacgtcacagcaacag 419
Db 323 AATCATCGGACCGAAGGCTTCTGGAGGCCCTTTCGAGGGCCCTCAACGCTCATGATCATGGG 382
Qy 420 cgcagggcctgccacgcctttatttgcctgtacgaaagttaaataaacacattgag 479
Db 383 TGCAGGGCCGGCCATGCCATGATATTTGGCTGCTATGAACACATGAAGAGGACTTTAAA 442
Qy 480 tgaatgaatcacccctggggccaatagccattatggcaatggtgcggccggtgtgtggc 539
Db 443 TGACGTTTTCCACCACCAAGGAACACACCACTAGCCACGGGATAGCTGGGAGTAGGC 502
Qy 540 aacattacttcattgatgcagccatgaaccctgcggaagtgttcgaagcagaggtatgcagat 599
```

```
Db 164 ATGATGGGACTGCCGGATGGCGGTGTCGGCAGCAAGGACGCGGTGCGGAGGACTAC 223
Qy 190 gagcgctgcgggtgagccactgtcaccaacacacatggtgagcagcgccgtgagcggg 249
Db 224 GAGAACTACCGACCGCGCTCGGTGTCACACACACATGACGCGGAGCGATGCCGGG 283
Qy 250 atcctgagcactcgtgatgtacccatgcactgcgtcaagaccgcgagtcagagacta 309
Db 284 ATTCTGAGCACTCCATCAATGATCCCGGTGACATCCGTGAAGACACGATGACAGTTG 343
Qy 310 cagcctgaccagctgcccctgcgaatggtgtgagggccctctgagagctataaga 369
Db 344 AATCCAGATCCAAAGCCCGGTATACAGCATCTATGCGCGCCCTCAAGAGGATCATGCAC 403
Qy 370 acgagggccctatgagggcccatgaggggctgaagctcacagacaacagcgagggcct 429
Db 404 ACTAAGGCTTCTGGAGGCGCCCTCGGGGCGCTGAACGTGATGATGGGTGCGAGGCCC 463
Qy 430 gccacgcccctttatttgcctgtacgaaagttaaataaacacattgagtgatgaatc 489
Db 464 GCGACGCCATGATTTTGGCTGTATGAACACATGAAGGACCTTAATGACGTTTTC 523
Qy 490 caccctggggcaatagccattatgccaatggtgcggccgggtgtgtgcaacattactt 549
Db 524 AGCCACCAAGGAACAGCCCATCTAGCTAATGGGTAGCTGGGAGTATGGCCACCCTACTC 583
Qy 550 catgatcagccatgaacccctgcgaagtgtcaagcagagagatgcagatgataactca 609
Db 584 CACGATCAGTAATGAATCAGCAGAGAGTGTGAACACGCGCTTACAGATGTACAACTCC 643
Qy 610 ccatcaccacgggtgcagagactgtgcagggcagtggtggcaaatgaagggccggggcc 669
Db 644 CAGCACGATCAGCCTTCAGTTGTATCCGGACAGTGTGGCGGACCGAGGGTTGGGGCC 703
Qy 670 tttaaccgagctacacccacccagctgaccatgaaccttcccttcccaagccattcattc 729
Db 704 TTCTACAGGAGTTACACACACAGCTGACCATGAATATCCCTTCCAGTCAATTCACCTC 763
Qy 730 atacattataattctgcagagcactttaacccccagagaggtacaacccaagctcc 789
Db 764 ATACCTATGAGTTCTGCAAGGAGCAAGTCAACCCCTCGCCGGGACTACACCCACAGCTC 823
Qy 790 cagctctctctgaggttcgcagaggagctgagctgcgcagccacacaccccaactggac 849
Db 824 CACATCATCTCAGGAGGCGCTGGCGGGGCACTGGCCGCACTGCCACCCCGCTGGAC 883
Qy 850 gtttgaaaaactgctcaacacccagaggtccttggttgaaact---cacacattaca 906
Db 884 GTCTGCAAAACCTCTCAACACGCAAGAGAACATGCTCTCTCCCTGGCCACGTCAGC 943
Qy 907 ggacatacacagcagtggttagtccttcagcaggtatataaagttagtggtggtgacc 966
Db 944 GGCCGGCTGTGGGCATGGCCATGCCCTTCGGACGGGTATACAGCTCAACGGCCTTGCC 1003
Qy 967 gctatttcgaggggtgcagggccagagtaattaccagatccccctccacagccatcgca 1026
Db 1004 GGCTATTTCAAGGATCCAGGCTCGAGTCACTTACCAGATGCTTCCACCGCATCTCC 1063
Qy 1027 tggctgtgatgagttcttcaataactaatcaataaaaggcaagaag 1077
Db 1064 TGGTCCGCTTTATGAGTCTTCAAGTACATCTTACAAAGAGGCGAGCTGGAG 1114
```

```

Db 503 CACCTGCTCCAGATCGGGAATGAATCCAGAGAAGTGTGAAGCAGCGCTTGCGAT 562
Qy 600 gtacaaactcaacattaccacccgggtgacagactgtgtacggcagtggtggcaaaatgaag 659
Db 563 GTAACTCGCAGCAGCGGTGAGCAATCAGTGTGATCGGACGGTGTGTGAGGACCGAGG 622
Qy 660 ggcggggccctttaccgacgtacacacccacccagctgacacatgaacgttcccttccaaagc 719
Db 623 GTTGGGGCCCTTACCGGAGCTACACCGAGCTACACCGAGCTGACCATGAACATCCCTTCCAGTC 682
Qy 720 catcaactaatgacctaatctctcagagacacactttaaccccccagagacggtacaa 779
Db 683 CATCCACTTTCATCACTATGATGATTCCTGAGGAGAGGTCAACCCCAACCGAGACTCAAA 742
Qy 780 cccaagctcccaagctccctctgagcttgccgagagagctgtagtgcgcgagcagcaaac 839
Db 743 CCGCAGTCCACATCATCTCAGCGGGGCTGCGCGGGCCCTGCCGGCGGCGCCACGAC 802
Qy 840 cccactgacgttttgcaaaacactgctcaacacccagagtgctcttggttt---gaactc 896
Db 803 CCCCTGGACGCTCTGAAGACCTTCTGAACACTCAGGAGAACGTGGCCCTCTCGCTGGC 862
Qy 897 acacattacagacatatcacagcgatgctagtgctctcagagcggtatatcaagttag 956
Db 863 CAACATCAGCGCGCGCTGTCGGGTATGGCCAATGCCCTTCGGAGCGGTGTACAGCTCAA 922
Qy 957 tgggtgacccgctatttccaggggtgagcagcagcagagtaattaccagatccccccac 1016
Db 923 CGGCTG-CCGGCTACTTCAAGGATCAGCGCGGTGTCTATCCACAGATGCCCTCCAC 981
Qy 1017 agccatcgatgctgtgtatgagttcttcaaatcacctaatcaactaaaagcca 1070
Db 982 CGCCATTTCTTGCTGTATGATGTTCTCAAGTACTTCTTCAACAAAGCGCCA 1035

RESULT 13
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
SCHOOL Juridical Person Nihon University (JP)
FEATURES
source
1. 483
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN

Query Match 38.2%; Score 417.8; DB 6; Length 483;
Best Local Similarity 93.2%; Pred. No. 5.5e-65;
Matches 437; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 462 gtaaaaaagacatigagtgtatcaatccctgggggcaatagccatttgcgaatgg 521
Db 15 GTTAAAAAGACATGTGAGTACGTAAATCCACCGGGGCAATAGCCATATTGCCAATGG 74
Qy 522 tgcggccgggtgtgtggcaacattacttoatgatgcagccatgaacccctgcggaagtgt 581
Db 75 TGCAGCCGGATGTGTGGGACATATTCTTATGATGAGCCATGAATCCAGCGGAAGTCT 134
Qy 582 caagcagagatgcagatgtacaactcaccataccacccgggtgacagactgtgtacgggc 641

```

```

Db 135 CAACGAGGATGAGATGTACAACCTCGCGTACCACCGCTGACAGACTGTGTTCGGGC 194
Qy 642 agtctgcaaaatgaagggccggggccctttttaccgagctacacccacagctgacct 701
Db 195 AGTGTGGCAAAATGAAGGGCGCGGGCCCTTTTACCGAGCTTACAGCCCCAGCTGACCAT 254
Qy 702 gaacgttctcttccaagccattcattcatgacctgaattcctcagcagacactttaa 761
Db 255 GAATGTCCCTTCCAGCCATTCACTTATGACCTATGAGTCTCTGCAAGAGACTTTAA 314
Qy 762 cccccagagcgggtacacacccaaagctccacagctcctctctgagcttgcgcagagctgt 821
Db 315 CCCCAGAGACGGTACACCCAGCTCCCATGTGCTCTGTGGGCTCGCGCAGGAGCAGT 374
Qy 822 agctgcgcgagcccaacaccccccactggagcttgcacaaacactgctcaacaccccgagctc 881
Db 375 AGCTGCCGCCGCCAACACCCACTGGAGCTTTGTAAAAACACTGTCTAAACACCCAGGAATC 434
Qy 882 cttggcttgaaactcaacacattacagacatatcacagcgaatgagctag 930
Db 435 CTGGCTTTGAACCTAAACATTACTGGACATCATCAGGCATGGCTAGT 483

RESULT 14
AX071440
LOCUS AX071440 401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 401)
Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lamson,G., Dmanac,R., Crkenjakov,R., Dmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
source
1. 401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature
1. 401
/note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 35.8%; Score 392; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 gtctacagcctgacccagctgcccctatcgcaatgtgttgaggccctctgagagatta 364
Db 9 GTCTACAGCCTGACCCAGCTGCCCTATCGCAATGTGTGGAGGCCCTCTGGAGGATTA 68
Qy 365 taagaacgagggcctatggagcccatgaggggctgaacgtcaacagcagcgacg 424
Db 69 TAAGAAGCAGGAGGCTATGGAGGCCCATGAGGGGCTGAACCTCAGACACAGCGGCAG 128
Qy 425 ggcctgccaccccttatttgcctgctacgaaaagttaaaaagacattgagtgcg 484
Db 129 GGCCTGCCACGCCCTTTATTTTGTCTGCTACGAAAGTTAAAAAGACATTGAGTGATG 188
Qy 485 taatcacccctggggcaaatagccattatgccaatgtgctggccgggtgtgtggcaaat 544
Db 189 TAATCCACCCCTGGGGCAATAGCCCATATTGCAATGGTGGCGCCGGGTGTGTGGCAACAT 248

```

```
Qy 545 tacttcattgatgcagccatgaaccctgcggaagtgttcaagcagagagtgatgcagatgtaca 604
|||||
Db 249 TACTTCATGATGCACCCATGAACCTCGCGAAGTGGTCAACGAGAGGATGCAGATGTACA 308
605 actcaccataccacccgggtgacagactgttacggcgagtggtgcacaaatgaaggggccg 664
|||||
Db 309 ACTCACCATACACCGCGGTGACAGACTGTGACGGCGAGTGTGGCAAAATGAAGGGCGG 368
665 gggccttttaccgagctacacccccagctg 696
|||||
Db 369 GGGCCTTTTACCGCAGCTACACACCCACCCAGCTG 400

RESULT 15
AF361699 5869 bp mRNA linear ROD 07-NOV-2001
LOCUS Mus musculus mitochondrial carrier-like protein mRNA, complete cds;
DEFINITION nuclear gene for mitochondrial product.
ACCESSION AF361699
VERSION AF361699.1 GI:16755527
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE A novel mouse mitochondrial carrier protein gene is up-regulated
from young to adult NOD mice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2001) Department of Pathology, Immunology and
Laboratory Sciences, University of Florida, 1600 SW Archer Road,
Room D6-15, Gainesville, FL 32610, USA
FEATURES
source
1. 5869
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="14"
/map="between D14Mit124 and D14Mit236"
116..664
/note="MCLP"
/codon_start=1
/evidence=not_experimental
/product="mitochondrial carrier-like protein"
/protein_id="AAL27990.1"
/db_xref="GI:16755528"
/translation="MELRRGGVGNQAGRMDGDCRDGGCGSKDAGSEYENLPISAS
VSTHMTAGAWAGLEHSIMIPVDSVKTRMOSLAPDPDKARYTSIYGALKRIMHTEGFWR
PLRLGNVMMGAGPAHAMFYACYNMKRTLNDVFSHOGNSHLANGILKAFVWSWEALL
SGASSFSPSNLHPROTENSRTV"
BASE COUNT 1387 a 1427 c 1390 g 1665 t
ORIGIN

Query Match 25.7%; Score 281; DB 10; Length 5869;
Best Local Similarity 70.1%; Pred. No. 1e-40;
Matches 393; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 520 ggtcggccgggtgtgtggcaaatcttcatgatgcagccatgaacccctgcggaagtg 579
|||||
Db 2311 GGGGTAGCTGGGAGTATGGCCACCCTACTCCAGCATGCAGTAATGAATCCAGCAGAAGTG 2370
580 gtcaagcagagatgcagatgtacaactcaccatcaccgggtgacagatgtgtaccg 639
|||||
Db 2371 GTGAACAGCGCTTACAGATGTACAATCCAGCACCAGTCAGTGTGTGTCGG 2430
940 gcagtgtggcaaatgaaggggcggggcccctttaccgcagctacacccaccgctgacc 699
```

```
Db 2431 ACAGTGTGGCGGACGAGGGGTTGGGGCTTCTACAGAGATTACACACACAGCTGACC 2490
Qy 700 atgaacgttcctttccaagccattcaacttcattgaactatgaattcctgcagagacactt 759
|||||
Db 2491 ATGAATATCCCTTCCAGTCAATTCACTTCACTATGAGTTTCTGCAGGAGCAAGTC 2550
760 aacccccagagagtgatacaacccaagctcccaacgtccctctctgagcttgcgcagagct 819
|||||
Db 2551 AACCTCGCGGGGACTACAACCCACAGTCTCACATCATCTCAGGAGGCTGCGCCGGGCA 2610
820 gtactgcgcgagccacacccccactggacgtttgcaaaacactgctcaacacccaggag 879
|||||
Db 2611 CTGGCCGCGAGCTGCCACCAACCCCGCTGGAGGCTCCTCAAAACCTCTCTCAACACGAGGAG 2670
880 tctctggctttgaac---tcacacattacaggagacatacacaggcatggttagtgccttc 936
Db 2671 AACATGGGCTCTCTCCCTGGCCCAACGTCAGCGGCCGGCTGTCGGGATGGCCAAATGCCCTTC 2730
937 aggcaggttatatcaagttagtggtgggtgaccgcctatttccgaggggtgcagggccagagta 996
Db 2731 CGGACGGTGTACCAGCTCAACGGCCTTGGCGGCTATTTCAAAGGCATCCAGGCTCGAGTC 2790
997 atttaccagatccctccacagccatcgcatggtgtgtgtatgatttcttcaaatcacta 1056
Db 2791 ATCTACCAGATGCCCTTCCACCGCCATCTCCTGGTCCGTTTATGAGTTCTTCAAGTACATC 2850
1057 atcactaaaaggcaagaagag 1077
Db 2851 CTTCAAAAGAGGCGAGCTGGAG 2871
```

Search completed: August 27, 2002, 02:34:02
Job time: 7232 sec

us-09-870-113-1.rge

Tue Aug 27 07:53:39 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:34:40 ; Search time 209.67 seconds

(without alignments)
8966.573 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atggaggttgagggcgggg.....agtggaggctgcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :			
N_Geneseq_032802.*			
1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987.4	90.2	1336	21	Human uncoupling p
2	972.4	88.8	1322	22	Human transport pr
3	958	87.5	1294	22	Human polynucleoti
4	916.5	83.7	1316	22	Human polynucleoti
5	872	79.6	1418	21	Human ORFX ORF398
6	639	58.4	710	20	Human gene express
7	576	52.6	2502	22	Human secreted pro
8	556	50.8	2037	24	cdNA sequence #574
9	520	47.5	997	22	Human mitochondria

10	508	46.4	27960	22	AAK69779	Human immune/haema
11	508	46.4	27960	22	AAK73320	Human immune/haema
12	434.2	39.7	1305	22	AAF5920	Human mitochondria
13	417.8	38.2	483	24	ABI98871	Mouse ischaemic co
14	392	35.8	401	22	AAF66156	Novel human polynu
15	380.8	34.8	1716	21	AACT77189	Human ORFX ORF2744
16	342.2	31.3	1913	22	AAH89944	Human bone marrow
17	338.4	30.9	1902	22	AAK52172	Human polynucleoti
18	330.2	29.2	2562	22	AAH90057	Human bone marrow
19	300	27.4	300	20	AAZ13026	Human gene express
20	277.2	25.3	1814	22	AAK94871	Human full-length
21	276	25.2	1015	22	AAK52501	Human mitochondria
22	276	25.2	1673	22	AAK53156	Human polynucleoti
23	276	25.2	2710	21	AAK66452	Human secreted pro
24	275.6	25.2	1890	22	AAK94897	Human full-length
25	272.2	24.9	600	22	AAH22163	Human mitochondria
26	250	22.8	8211	22	AAK80490	Human immune/haema
27	181	16.5	339	22	AAK58672	Human immune/haema
28	173.6	15.9	842	22	AAK92357	Human cdna 5'-end
29	173.6	15.9	842	22	AAK93895	Human cdna clone r
30	172.6	15.8	1468	21	AACT77175	Human ORFX ORF2730
31	172	15.7	3028	22	AAK91320	Human polynucleoti
32	171.6	15.7	2137	23	ABL08627	Drosophila melanog
33	170.4	15.6	622	22	AAK91764	Human cdna 5'-end
34	170.4	15.6	622	22	AAK93200	Human cdna clone r
35	164.2	15.0	792	21	AACT77177	Human ORFX ORF2732
36	161.4	14.7	1548	23	AAK55357	DNA encoding novel
37	160	14.6	1758	20	AAV84573	Human secreted pro
38	160	14.6	1758	22	ABA83356	Human secreted pro
39	159.6	14.6	504	23	AAK55354	DNA encoding novel
40	153.6	14.0	338	22	AAK23288	Human prostate can
41	138.2	12.6	975	21	AACT77173	Human ORFX ORF2728
42	113.4	10.4	599	22	AAK24885	Human ovarian PCR
43	113.4	10.4	599	22	AAH83525	Human ovarian tumo
44	109.8	10.0	2837	22	AAK80493	Human immune/haema
45	92	8.4	5150	18	AAK68952	Blackcurrant fruit

ALIGNMENTS

RESULT 1	
AAK90457	
ID	AAK90457 standard; cdna; 1336 BP.
XX	AAK90457;
AC	AAK90457;
XX	
DT	12-MAR-2001 (first entry)
DE	Human uncoupling protein cdna #6.
XX	
KW	Human; uncoupling protein; immunosuppressive; antiarthritic;
KW	antirheumatic; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynaeological; vulnery; thrombolytic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
XX	infertility; ss.
OS	Homo sapiens.
XX	
PN	W0200061614-A2.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09534.
XX	
PR	09-APR-1999; 99US-0128701.
PR	08-JUL-1999; 99US-0142821.
PR	18-AUG-1999; 99US-0149448.
PR	12-NOV-1999; 99US-0164751.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	

DE	Human polynucleotide SEQ ID NO 1078.
XX	
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
XX	WPI; 2001-442253/47.
DR	P-PSDB; AAM39719.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Claim 1; SEQ ID NO 1078; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;

Query Match	83.7%	Score 916.6;	DB 22;	Length 1316;
Best Local Similarity	97.9%	Pred. No. 1.8e-195;		
Matches 939: Conservative	0;	Mismatches 19;	Indels 1;	Gaps 1;

Qy	138	cgggcctcagggcccccggtacacaagaatccggactccggcccgactacgaggcgct	197
Dy	138	cgggcctcagggcccccggtacacaagaatccggactccggcccgactacgaggcgct	197
Db	15	cgagcctcgagggcccccggtacatcagatccggactccggcccgactacgaggcgct	74
Qy	198	gccggctggagccactgtcacacgcacat-gtggcaggccgctggcaggatcctgg	256
Dy	198	gccggctggagccactgtcacacgcacat-gtggcaggccgctggcaggatcctgg	256
Db	75	gtcggcctggagccactgtcacacgcacatgtggggtatgcccagagcattggtctgg	134
Qy	257	agcactcgtatatacccatcactcgtgtcaaaccccgatcacagatctacagctg	316
Dy	257	agcactcgtatatacccatcactcgtgtcaaaccccgatcacagatctacagctg	316
Db	257	agcactcgtatatacccatcactcgtgtcaaaccccgatcacagatctacagctg	316

RESUL
AAI58875XX
XX
XX

XX

XX

Db 135 agcactgctgtagtaccctcagctcgtagtaccgagcagtagtcatgctacagtctg 194
Qy 317 accaagctgcgcgtatgcgaatgtgttgaggccctctgaggattataagaacggagg 376
Db 195 accagctggcgcgtatgcgaatgtgttgaggccctctgaggattataagaacggagg 254
Qy 377 gcctatgagcccatgaggggctggaacgtcacagcaacagcagcagcagcagcagc 436
Db 255 gcctatgagcccatgaggggctggaacgtcacagcaacagcagcagcagcagcagc 314
Qy 437 ccctttatttgcctgctacgaaagttaaaagacattgagtgatgtaattccacctg 496
Db 315 ccctttatttgcctgctacgaaagttaaaagacattgagtgatgtaattccacctg 374
Qy 497 ggggcaatagccattatgcgaatggtgcgcgcgggtgtgtggaacattacttcatg 556
Db 375 ggggcaatagccattatgcgaatggtgcgcgcgggtgtgtggaacattacttcatg 434
Qy 557 cagccatgaacccctgcggaagtgttcaagcagagagatgcagatgtacaaactcaccatacc 616
Db 435 cagccatgaacccctgcggaagtgttcaagcagagagatgcagatgtacaaactcaccatacc 494
Qy 617 accgggtgacagactgtgtacggcagtggtgcaaaatgaaggggccgggcccctttacc 676
Db 495 accgggtgacagactgtgtacggcagtggtgcaaaatgaaggggccgggcccctttacc 554
Qy 677 gcactcacaccagcagctgacatgaactgaactgtcttccaaagcattcaattcatgacct 736
Db 555 gcagctacaccagcagctgacatgaactgaactgtcttccaaagcattcaattcatgacct 614
Qy 737 atgaattctgcagagacacttcaaccccgagagcgtgtacaaacccaagctcccaagctcc 796
Db 615 atgaattctgcagagacacttcaaccccgagagcgtgtacaaacccaagctcccaagctcc 674
Qy 797 tctctgagcttgcgcaggagctgtagtgcgcagccacacacccactggagctttgca 856
Db 675 tctctgagcttgcgcaggagctgtagtgcgcagccacacacccactggagctttgca 734
Qy 857 aaacactgtcaacaccagagctccttgcttggaactcacattacagagacatatca 916
Db 735 aaacactgtcaacaccagagctccttgcttggaactcacattacagagacatatca 794
Qy 917 caggcatggctagtgcttcaggagcgttatcaagtagtggtgggtgaccgctatttcc 976
Db 795 caggcatggctagtgcttcaggagcgttatcaagtagtggtgggtgaccgctatttcc 854
Qy 977 gaggggtgcagccagagtaattaccagatccctccacagccatgcattggtctgt 1036
Db 855 gaggggtgcagccagagtaattaccagatccctccacagccatgcattggtctgt 914
Qy 1037 atgattcttcaatacctaatcaataaaggcaagagagtaggagcgtggaagtga 1095
Db 915 atgattcttcaatacctaatcaataaaggcaagagagtaggagcgtggaagtga 973

RESULT 5
AAC74843
ID AAC74843 standard; cDNA; 1418 BP.
AC AAC74843;
XX
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF398 polynucleotide sequence SEQ ID NO:795.
XX
KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX Homo sapiens.
OS
XX WO200058473-A2.
PN
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
PF
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shinkets RA, Leach M;
FI
XX WPI: 2000-602362/57.
DR P-PSDB; ABA40634.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 853-854; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;
SQ

Query Match 79.6%; Score 872; DB 21; Length 1418;
Best Local Similarity 91.3%; Pred. No. 1.7e-185;
Matches 975; Conservative 0; Mismatches 0; Indels 93; Gaps 1;
Qy 121 gccggcgccggggaggccggcctgagcccccggagcccggtacgacagatccggactccggc 180
Db 1 gccggcgccggggaggccggcctgagcccccggagcccggtacgacagatccggactccggc 60
Qy 181 ccgactacagggcgctccggcctggagccactgtccacgcacatggtgcagcgcc 240
Db 61 ccgactacagggcgctccggcctggagccactgtccacgcacatggtgcagcgcc 120
Qy 241 gtggcagggatccctggagcactggtgatgtaccccatcgactcgctcaagaccggatg 300

Db 166 ggtgtcaagcagagtgatgagatgtacaaactcaaccataccaccgggtgacagactgtgta 225
Qy 637 cggcagtgtagcaaaatgaaggccggcgccctttttacacagactacacaccagctg 696
Db 226 cggcagtgtagcaaaatgaaggccggcgccctttttacacagactacacaccagctg 285
Qy 697 accatgaacgttctcttccaagccattcaacttcacatgacattgaattcctgcagagcac 756
Db 286 accatgaacgttctcttccaagccattcaacttcacatgacattgaattcctgcagagcac 345
Qy 757 ttttaacccccagagcaggtacaaaccccaagctccacagctctctctgagcttgccagga 816
Db 346 ttttaacccccagagcaggtacaaaccccaagctccacagctctctctgagcttgccagga 405
Qy 817 gctagtctgcgcagcacaacccccactgacgttttgcaaaacactgctcaacaccag 876
Db 406 gctagtctgcgcagcacaacccccactgacgttttgcaaaacactgctcaacaccag 465
Qy 877 gagtcttggtttgaactcacattacagagcatatcacagcagcatggctagtgccttc 936
Db 466 gagtcttggtttgaactcacattacagagcatatcacagcagcatggctagtgccttc 525
Qy 937 aggcaggtatatcaagttagtggtgggtgacccctatttccaggggtgcagggcagagta 996
Db 526 aggcaggtatatcaagttagtggtgggtgacccctatttccaggggtgcagggcagagta 585
Qy 997 atttaccagatccctccacagccatcgcatggtctgtgtatgattcttcaataccta 1056
Db 586 atttaccagatccctccacagccatcgcatggtctgtgtatgattcttcaataccta 645
Qy 1057 atcactaaaggcgaagagtgagggtgagcgttgccagtgga 1095
Db 646 atcactaaaggcgaagagtgagggtgagcgttgccagtgga 684

RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
AC AAS03906;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human secreted protein gene #25.
XX
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection; ss;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
OS Homo sapiens.
XX
PN WO200123598-A1.
XX
PD 05-APR-2001.
XX
PF 26-SEP-2000; 2000WO-US26324.
XX
PR 27-SEP-1999; 99US-0155807.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis G, Ruben SM, Rosen CA;
XX
DR WPI; 2001-281684/29.
DR P-PSDB; AAU01950, AAU01989.
XX
PT Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Disclosure; Page 460-461; 518pp; English.

XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
CC PCR primers of the invention. acid of the invention. Secreted proteins and
CC their related nucleic acids can be used in the diagnosis of or
CC susceptibility to a pathological condition by determining the presence or
CC absence of a mutation in a nucleic acid or the presence or amount of
CC expression of a secreted protein. The sequences are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
CC polypeptides can also be used in alleviating symptoms associated with
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). The disorders include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The peptides can also be used to aid wound healing and epithelial cell
CC proliferation, to help prevent skin ageing due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues, in chemotaxis and
CC as a food additive or preservative to alter storage capabilities.
XX
SQ Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;

Query Match 52.6%; Score 576; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 2.5e-119;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 520 ggtgcgcgcgggtgtgtggcaacattacttcatgatgcagccatgaacctgcggaagtg 579
Db 1513 ggtgcgcgcgggtgtgtggcaacattacttcatgatgcagccatgaacctgcggaagtg 1572
Qy 580 gtcaagcagagagtgacagatgtacaaactcaccacccgggtgacagactgtgtacgg 639
Db 1573 gtcaagcagagagtgacagatgtacaaactcaccacccgggtgacagactgtgtacgg 1632
Qy 640 gcagtgtggcaaaatgaaggccggcgcccttttaccgcagctacacaccaccagctgacc 699
Db 1633 gcagtgtggcaaaatgaaggccggcgcccttttaccgcagctacacaccaccagctgacc 1692
Qy 700 atgaacgttctcttccaaagccattcaacttcatgacattgaattcctgcagagcaacttt 759
Db 1693 atgaacgttctcttccaaagccattcaacttcatgacattgaattcctgcagagcaacttt 1752
Qy 760 aacccccagagagtgacacaccccaagctccacagctcctctctgagcttgccagagact 819
Db 1753 aacccccagagagtgacacaccccaagctccacagctcctctctgagcttgccagagact 1812
Qy 820 gtactgcgcgcagcacaaccccaactgacgttttgcaaaacactgctcaacaccagag 879
Db 1813 gtactgcgcgcagcacaaccccaactgacgttttgcaaaacactgctcaacaccagag 1872
Qy 880 tccttggtctttgaactcacacattacagggacatatacagagcattggttagtgccttcagg 939
Db 1873 tccttggtctttgaactcacacattacagggacatatacagagcattggttagtgccttcagg 1932
Qy 940 acggtatatcaagtagtggtgggtgaccgcctatttccagaggggtgcagggccagagtaatt 999
Db 1933 acggtatatcaagtagtggtgggtgaccgcctatttccagaggggtgcagggccagagtaatt 1992
Qy 1000 taccagatccctccacagccatcgatggtctgtgtatgattcttcaatacctaacc 1059
Db 1993 taccagatccctccacagccatcgatggtctgtgtatgattcttcaatacctaacc 2052
Qy 1060 actaaagggcaagagtagtgagggtggcgaagtga 1095
Db 2053 actaaagggcaagagtagtgagggtggcgaagtga 2088

RESULT 8
AAS62787
ID AAS62787 standard; cDNA; 2037 BP.

XX AAS62787;
AC
XX
DT 14-FEB-2002 (first entry)
DE
XX cDNA sequence #574 encoding novel human secreted protein.
XX
XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200177291-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 29-MAR-2001; 2001WO-US10485.
XX
XX 06-APR-2000; 2000US-195604P.
PR
XX
XX (GENY) GENETICS INST INC.
PA
XX
XX Wong GG, Clark HF, Fichtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
PI
XX
XX WPI; 2002-010900/01.
DR
XX
XX New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
PT
XX
XX
PS Claim 1; Page 364-365; 391pp; English.
XX
XX The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
XX Sequence 2037 BP; 475 A; 522 C; 524 G; 515 T; 1 other;
SQ

Query Match 50.8%; Score 556; DB 24; Length 2037;
Best Local Similarity 99.0%; Pred. No. 7e-115;
Matches 570; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 520 ggtcgccggtgtgtggaacattactatgatgagccatgaacccctcggaagt 579
Db 1138 ggtcgccggtgtgtggaacattactatgatgagccatgaacccctcggaagt 1197
QY 580 gtcaagcagagatgcagatgtacaactaccatccacccgggtgacagactgtacgg 639
Db 1198 gtcaagcagagatgcagatgtacaactaccatccacccgggtgacagactgtacgg 1257
QY 640 gcagttgtgcaaaatgaagggccggggcctttaccgcagctacaccaccagctgacc 699
Db 1258 gcagttgtgcaaaatgaagggccggggcctttaccgcagctacaccaccagctgacc 1317
QY 700 atgaacgttcctttcccaagccaattcaattcagctatgaattcctcagagacattt 759
Db 1318 atgaacgttcctttcccaagccaattcaattcagctatgaattcctcagagacattt 1376

QY 760 aacccccagagcgtacaaacccaagctcccacgtctctctgtgagcttgcgagagct 819
Db 1377 aacccccagagcgtacaaacccaagctcccacgtctctctgtgagcgtgcgagagct 1436
QY 820 gtactgcgcagcagcaacccccactgacgtttgcaaaacactgctcaacacccaggag 879
Db 1437 gtactgcgcagcagcacatccccactggcgtttgcaaaacactgctcaacacccaggag 1496
QY 880 tcttggctttgaactcacacattacaggacatatcaggagcataggctagtccttcagg 939
Db 1497 tcttggctttgaactcacacattacaggacatatcaggagcataggctagtccttcagg 1556
QY 940 acggtatatcaagtgtgggtgaccgctatttccagaggggtgcagccagagtaatt 999
Db 1557 acggtatatcaagtgtgggtgaccgctatttccagaggggtgcagccagagtaatt 1616
QY 1000 taccagatccccctccacagccatcgatgctgtgtgtatgagttcttcaaaatacctaattc 1059
Db 1617 taccagatccccctccacagccatcgatgctgtgtgtatgagttcttcaaaatacctaattc 1676
QY 1060 actaaaggccaagaagatggagggtgccaagtga 1095
Db 1677 actaaaggccaagaagatggagggtgccaagtga 1712
RESULT 9
AAH22162
ID AAH22162 standard; cDNA; 997 BP.
XX
AC AAH22162;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human mitochondrial solute carrier (hMSC-homologue) encoding cDNA SEQ.6.
XX
KW Human; hMSC; hMSC-homologue; mitochondrial solute carrier; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 379..909
FT /*tag= a
FT /product= "mitochondrial solute carrier (hMSC-homologue)"
XX
PN CN1281895-A.
XX
PD 31-JAN-2001.
XX
PF 27-JUN-2000; 2000CN-0116795.
XX
PR 27-JUN-2000; 2000CN-0116795.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Li N, Xu X, Xiao H;
XX
DR WPI; 2001-282781/30.
XX
XX P-PSDB; AAB98221.
PT New human mitochondrial solute carrier (hMSC) protein isomer and its
PT code sequence -
XX
PS Claim 1; Page 17-18 (disclosure); 22pp; Chinese.
XX
CC The present invention describes a human mitochondrial solute carrier
CC protein, designated hMSC-homologue, which is expressed in human
CC hypophysis. Also described are methods for the preparation and
CC detection of hMSC-homologue protein and nucleotide sequences. The
CC present sequence encodes hMSC-homologue, as given in the present
CC invention.
XX
XX Sequence 997 BP; 240 A; 261 C; 242 G; 254 T; 0 other;
SQ

Query Match		47.5%;	Score 520;	DB 22;	Length 997;
Best Local Similarity		100.0%;	Pred. No. 6.7e-107;		
Matches	520;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	576	agtggtcaagcagagatgcagatgtacaactaccataccaccgggtgacagactgtgt	635		
Db	390	agtggtcaagcagagatgcagatgtacaactaccataccaccgggtgacagactgtgt	449		
QY	636	acgggcagtggtgcaaaatgaaggccggcggccttttacgcagctacacacccagct	695		
Db	450	acgggcagtggtgcaaaatgaaggccggcggccttttacgcagctacacacccagct	509		
QY	696	gaccatgaacgtctcttccaaagcattcactctcatgacctatgaattccctgcaggagca	755		
Db	510	gaccatgaacgtctcttccaaagcattcactctcatgacctatgaattccctgcaggagca	569		
QY	756	ctttaaccccccagagcaggtacaaccccaagctcccaagcttctctgtgagcttgcgcagg	815		
Db	570	ctttaaccccccagagcaggtacaaccccaagctcccaagcttctctgtgagcttgcgcagg	629		
QY	816	agctgtagctccgcagcagcaaccccaactggagcgtttgcaaaacactgctcaacaccca	875		
Db	630	agctgtagctccgcagcagcaaccccaactggagcgtttgcaaaacactgctcaacaccca	689		
QY	876	ggagtccttgcttgtaactcacattacagacatatacacagcatggtcgtagtgcctt	935		
Db	690	ggagtccttgcttgtaactcacattacagacatatacacagcatggtcgtagtgcctt	749		
QY	936	caggacggtatatcaagttagtggtggtagccgcctatttccgaggggtgcaggccagagt	995		
Db	750	caggacggtatatcaagttagtggtggtagccgcctatttccgaggggtgcaggccagagt	809		
QY	996	aatttacagatcccctccacagccatcgcatgctgtgtatgatttcttcaataacct	1055		
Db	810	aatttacagatcccctccacagccatcgcatgctgtgtatgatttcttcaataacct	869		
QY	1056	aatcactaaaagcaagaagtagtgaggctggcgaagtga	1095		
Db	870	aatcactaaaagcaagaagtagtgaggctggcgaagtga	909		
RESULT 10					
AAK69779					
ID	AAK69779	standard;	DNA;	27960	BP.
XX	AC	AAK69779;			
XX	DT	06-NOV-2001 (first entry)			
XX	DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.			
XX	KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;			
XX	KW	cytostatic; gene therapy; vaccine; metastasis; ds.			
XX	OS	Homo sapiens.			
XX	PN	W0200157182-A2.			
XX	PD	09-AUG-2001.			
XX	PF	17-JAN-2001; 2001WO-US01354.			
XX	PR	31-JAN-2000; 2000US-0179065.			
XX	PR	04-FEB-2000; 2000US-0180628.			
XX	PR	24-FEB-2000; 2000US-0184664.			
XX	PR	02-MAR-2000; 2000US-0186350.			
XX	PR	16-MAR-2000; 2000US-0189874.			
XX	PR	17-MAR-2000; 2000US-0190076.			
XX	PR	18-APR-2000; 2000US-0198123.			
XX	PR	19-MAY-2000; 2000US-0205515.			
XX	PR	07-JUN-2000; 2000US-0209467.			

PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.

PD 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Disclosure; SEQ ID NO 28132; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytotstatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 46.4%; Score 508; DB 22; Length 27960;
Best Local Similarity 99.8%; Pred. No. 6.3e-104;
Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 576 agtgtaacagcagaggatgcagatgtcaactaccataccacccgggtgcagactgtgt 635
Db 20845 agtgtaacagcagaggatgcagatgtcaactaccataccacccgggtgcagactgtgt 20904
QY 636 acgggcagtggtggcaaaatgaaggccggcgctttaccgcagctacacacccagct 695
Db 20905 acgggcagtggtggcaaaatgaaggccggcg-gctttaccgcagctacacacccagct 20963
QY 696 gaccatgaacgttcttccaagccattcacttcattcatgcctatgacctatgaattctgcaggagca 755
Db 20964 gaccatgaacgttcttccaagccattcacttcattcatgcctatgacctatgaattctgcaggagca 21023
QY 756 ctttaacccccagagcgggtacaaacccaagctccacgctctctctgagcttgcgagg 815
Db 21024 ctttaacccccagagcgggtacaaacccaagctccacgctctctctgagcttgcgagg 21083
QY 816 agctgtagtcgcccagcccaacacccactggacgttttgcaaaacactgctcaacaccca 875
Db 21084 agctgtagtcgcccagcccaacacccactggacgttttgcaaaacactgctcaacaccca 21143
QY 876 ggaagtccttggttgtaactcaacatatacaggacatatcacaggcagcatggttagtgcctt 935
Db 21144 ggaagtccttggttgtaactcaacatatacaggacatatcacaggcagcatggttagtgcctt 21203
QY 936 cagcacggtatatcaagttagtggtgacgcctatttccgaggggtgcagccagagt 995
Db 21204 cagcacggtatatcaagttagtggtgacgcctatttccgaggggtgcagccagagt 21263
QY 996 aatttacagatccctccacagccatcgcatggtctgtgtatgagtcttccaatacct 1055
Db 21264 aatttacagatccctccacagccatcgcatggtctgtgtatgagtcttccaatacct 21323
QY 1056 aataactaaaggcaagaagtggaagggtgcgaagtga 1095
Db 21324 aataactaaaggcaagaagtggaagggtgcgaagtga 21363

RESULT 12

AAF59920

ID AAF59920 standard; cDNA; 1305 BP.

XX

AC AAF59920;
XX
DT 04-MAY-2001 (first entry)
XX
DE Human mitochondrial solute carrier protein hMSC-o cDNA.
XX
KW Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
KW preparation; detection; ss.
XX
XX Homo sapiens.
XX
XX CN1269409-A.
XX
XX 11-OCT-2000.
XX
XX 17-MAR-2000; 2000CN-0114958.
XX
XX 17-MAR-2000; 2000CN-0114958.
XX
XX (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX
XX Zhang X, Gao X, Xiao H;
XX WPI; 2001-050544/07.
XX P-PSDB; AAB60658.
XX
XX New human mitochondrion solute carrier protein and its nucleic acid -
XX
XX Claim 1; Page 19-20; 21pp; Chinese.
XX
XX The invention relates to a novel human mitochondrial solute carrier
XX protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
XX expressed in normal human hypothalamus tissue. The invention also relates
XX to the preparation of hMSC-o proteins and nucleic acids, and the
XX detection of hMSC-o proteins and nucleic acids in a sample. The present
XX sequence represents cDNA encoding hMSC-o.
XX
XX Sequence 1305 BP; 292 A; 388 C; 363 G; 262 T; 0 other;
XX
Query Match 39.7%; Score 434.2; DB 22; Length 1305;
Best Local Similarity 69.5%; Pred. No. 1e-87;
Matches 619; Conservative 0; Mismatches 268; Indels 4; Gaps 2;
QY 183 ggactacagagcgctgcggctggagccactgtccacacgcacatggtggcagcgccgt 242
Db 57 ggactacagaaactgcgcactagcctcctgtccaccacacatgacagcagagcgat 116
QY 243 ggacgggactcgtgagcactgcgtgtaccccatgcactgcgtcaagaccccgatgca 302
Db 117 ggcgggactcgtggagcactgcgtgtaccccggtggaactcgtggaagacacgaatgca 176
QY 303 gagtctacagcctgacccagctgcccgctatcgcaatgtgttggaggccctctgaggat 362
Db 177 gaggttgagtcacagatcccaagcccgagtcacacagtatctacggagccctcaagaaat 236
QY 363 tataagaacagaggcgctatgagggcccatgagggggtgagcctcagacacacagcgcc 422
Db 237 catgagccaggaaggtctctggaggcccttgcgagggcgtcaacgctcatgatctatggtgc 296
QY 423 agggcctgccacgcccctttatttgcctgctacgaaaagttaaaaaagacatgagtga 482
Db 297 agggcggcccatgccaatgtatttgcctgtatgaaaacatgaaaaggactttaatga 356
QY 483 tgtaatccaccctcgggggcaatagccatatgtgcaatggtgcggccgggtgtgtggcaac 542
Db 357 cgtttccaccacaaaggaaacagccacccctagccacgggagtagctgggagtagtgccac 416
QY 543 attactcatgatgcagccatgaaccctgcggaagtgttcaagcagaggtgcagatgta 602
Db 417 cctgctccacgagcggtaatgaatgaatgcagagagtggtgaagcagcgctgcagatgta 476
QY 603 caactcaccataccacccgggtgacagactgtgtacgggagctgtggcaaaatgaaggggc 662


```

XX (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI: 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 818; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      35.8%; Score 392; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.2e-78;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 gtctacagcctgaccagctgcccgctatgcgaatgtgttgaggccctctggaggatta 364
Db 9 gtctacagcctgaccagctgcccgctatgcgaatgtgttgaggccctctggaggatta 68
QY 365 taagaacggaggccctatgagcccataggggctgaacgtcacagcaacaggcgacg 424
Db 69 taagaacggaggccctatgagggcccataggggctgaacgtcacagcaacaggcgacg 128
QY 425 ggccgtgccacgccccttattttgctgctacgaaaagttaaaagacattgagtgatg 484
Db 129 ggccgtgccacgccccttattttgctgctacgaaaagttaaaagacattgagtgatg 188
QY 485 taatcacccctggggcaatagcattatgcgaatggtgcggccggtgtgtggcaacat 544
Db 189 taatcacccctggggcaatagcattatgcgaatggtgcggccggtgtgtggcaacat 248
QY 545 tacttcatgatgcagcctgaacccctgcggaagtgttcaagcagagatgcagatgtaca 604
Db 249 tacttcatgatgcagcctgaacccctgcggaagtgttcaagcagagatgcagatgtaca 308
QY 605 actcacataccaccgggtgacagatgtgtgtacggcgagtggtggcaaatgaagggcccg 664
Db 309 actcacataccaccgggtgacagatgtgtgtacggcgagtggtggcaaatgaagggcccg 368
QY 665 gggccctttaccagcactacacaccagctg 696
Db 369 gggccctttaccagcactacacaccagctg 400

RESULT 15
AAC77189
ID AAC77189 standard; cDNA; 1716 BP.
XX

```

```

AC AAC77189;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF744 polynucleotide sequence SEQ ID NO:5487.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineuritic;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
OS
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX
XX P-PSDB; AAB42980.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 4661-4662; 5507pp; English.
XX
XX AAC7446 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antineuritic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;

```

Query Match 34.8%; Score 380.8; DB 21; Length 1716;

	Best Local Similarity	68.8%;	Pred. NO. 9.4e-76;			
	Matches	539;	Conservative	0;	Mismatches	242;
					Indels	3;
					Gaps	1;
Qy	230	agaccggatgcagaggtctacagctgaccagctgcccgcctatcgcaatgtgttgagg	349			
Db	248	agacacgaatgcagagtttgatccagatcccaagcccgatcacacagttctacggag	307			
Qy	330	ccctctggagattataagaacggaggccctatggaggcccatgaggggctgaacgtca	409			
Db	308	ccctcaagaataatgcggaccgaagctcttgaggcccttgagagcgtcaacgtca	367			
Qy	410	cagcaacggcgagggcctgccacgccctttatttgcctgctacgaaagtcaaaa	469			
Db	368	tgatcatgggtgcaggggccagcccatgccatgtatttgcctgctatgaaacatgaaa	427			
Qy	470	agacattgagtgtataatccacccctggggccaatggccattatggccaatggctggccg	529			
Db	428	ggactttaatgagcttttccaccaccaaggaacagccacctagccaacgggatagctg	487			
Qy	530	ggctgtgtgcaacattacttatgatgcagccatgaacccctgcggaagtgtcaacgaga	589			
Db	488	ggagtatggccacccctgtccacagtgcgttaatgaaatccacgagaagtgtgaagcgc	547			
Qy	590	ggatgcagatgtacaactcacatcacccgggtgacagactgtgtacggcagctgtgac	649			
Db	548	gctgcagatgtacaactgcagaccggtcagcaatcagctgtcatccgacggtgttggg	607			
Qy	650	aaatgaaggggcgggccctttaccgcagctcacaccaccagctgaccatgaacgttc	709			
Db	608	ggacggagggttggggcccttaccggagctacaccacgcagctgaccatgaacatcc	667			
Qy	710	cttccaaagccattcaacttcctgacctatgaattctcctcaggagcaactttaacccccaga	769			
Db	668	cctctcagctccatccacttcactatgactatgattctcctcaggagcaggtcaacccccacc	727			
Qy	770	gacggtaacacccaagctcccacgtccctctctgagcttgcgcagagctgactgcgcg	829			
Db	728	ggacctcaaccccagctccacatcatctcaggcggtgcggggccctcgcgcggg	787			
Qy	830	cagccacaacccccactggacgttttgcaaaaacactgtcaacaccacagagctccttggc	887			
Db	788	ccgcacgaacccccctggagctctgtaagacccttctgaacactcaggagaacgtggccc	847			
Qy	888	-tttgaactcaacattacaggagacatcacaggcatgctgtagtgccttcaggacggtat	946			
Db	848	tctcgctggcacaactcagcgccgctgtgcggtatggccaatgccttcgcggagcgtgt	907			
Qy	947	atcaagtgtgggggtgaccgcctatttccgaggggtgcagccagagtaattaccaga	1006			
Db	908	accagctcaacggcctggcgcggtacttcaaggcatccagggcggtgctatctaccaga	967			
Qy	1007	tccctccacagccatgcagctggtctgtgatgagttcttcaaataccataatcaataaa	1066			
Db	968	tgcctccacggccatttcttgctgtctgtatgattcttcaagtacttctcaccagc	1027			
Qy	1067	ggca	1070			
Db	1028	gcca	1031			

Search completed: August 27, 2002, 02:35:07
Job time: 7227 sec

THIS PAGE BLANK (USPTO)

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 01:24:15 ; Search time 49.51 seconds
(without alignments)
5432.613 Million cell updates/sec

Title: US-09-870-113-1
Perfect score: 1095
Sequence: 1 atggaggtggggggggggg.....agtggaggctggcaagtga 1095

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	8.4	5150	US-09-068-140A-14	Sequence 14, Appl
2	90.4	8.3	1311	US-09-068-140A-9	Sequence 9, Appl
3	62.4	5.7	319	US-09-165-264-8	Sequence 8, Appl
4	58.2	5.3	318	US-09-165-264-12	Sequence 12, Appl
5	58.2	5.3	320	US-09-165-264-14	Sequence 14, Appl
6	58.2	5.3	152331	US-09-128-155-16	Sequence 16, Appl
7	57.8	5.3	320	US-09-165-264-7	Sequence 7, Appl
8	57.6	5.3	320	US-09-165-264-13	Sequence 13, Appl
9	57.4	5.2	320	US-09-165-264-11	Sequence 11, Appl
10	50.6	4.6	833	US-08-403-852B-3	Sequence 3, Appl
11	50.6	4.6	833	US-08-510-646B-3	Sequence 3, Appl
12	50.6	4.6	833	US-09-231-818-3	Sequence 3, Appl
13	50.6	4.6	5392	US-08-403-852B-1	Sequence 1, Appl
14	50.6	4.6	5392	US-08-510-646B-1	Sequence 1, Appl
15	50.6	4.6	5392	US-09-231-818-1	Sequence 1, Appl
16	49.4	4.5	43280	US-08-804-227C-1	Sequence 1, Appl
17	47	4.3	1227	US-08-903-800A-1	Sequence 1, Appl
18	46.8	4.3	1931	US-09-130-114-2	Sequence 2, Appl
19	46.8	4.3	4257	US-08-690-473-1	Sequence 1, Appl
20	46.8	4.3	4257	US-09-259-821A-1	Sequence 1, Appl
21	46.8	4.3	4257	US-08-843-659-1	Sequence 1, Appl
22	46.8	4.3	12001	US-08-458-568A-11	Sequence 11, Appl
23	46.6	4.3	4403765	US-09-103-840A-2	Sequence 2, Appl
24	46.6	4.3	4411529	US-09-103-840A-1	Sequence 1, Appl
25	46.2	4.2	2823	US-08-398-008A-1	Sequence 1, Appl
26	46.2	4.2	2823	US-08-893-333-1	Sequence 1, Appl
27	46.2	4.2	3415	US-08-054-077C-1	Sequence 1, Appl

28	45.8	4.2	460	US-08-903-800A-4	Sequence 4, Appl
29	45.6	4.2	1300	US-08-483-533-39	Sequence 39, Appl
30	45.6	4.2	1300	US-09-283-471A-39	Sequence 39, Appl
31	45.6	4.2	53526	US-08-658-136-2	Sequence 2, Appl
32	45.6	4.2	53577	US-08-658-136-1	Sequence 1, Appl
33	45.4	4.1	1327	US-08-483-533-36	Sequence 36, Appl
34	45.4	4.1	1327	US-09-283-471A-36	Sequence 36, Appl
35	45.4	4.1	1335	PCT-US91-065332-1	Sequence 1, Appl
36	45.2	4.1	1028	US-08-118-200-1	Sequence 1, Appl
37	45.2	4.1	1028	US-08-458-745-1	Sequence 1, Appl
38	45.2	4.1	12001	US-08-458-568A-11	Sequence 11, Appl
39	45	4.1	8438	US-07-945-283-1	Sequence 1, Appl
40	44.6	4.1	5228	US-09-428-711A-15	Sequence 15, Appl
41	44.6	4.1	23673	US-09-773-816-1	Sequence 1, Appl
42	44.4	4.1	405	US-08-903-800A-3	Sequence 3, Appl
43	44.2	4.0	700	US-09-236-097-7	Sequence 7, Appl
44	44.2	4.0	1026	US-07-751-891B-24	Sequence 24, Appl
45	44.2	4.0	1868	US-08-658-883B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-068-140A-14

Sequence 14, Application US/09068140A

Patent No. 6281409

GENERAL INFORMATION:

APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor

APPLICANT: and Rex Michael Brennan

TITLE OF INVENTION: Blackcurrant Promoters and Genes

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068,140A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP96/04807

FILING DATE: No. 6281409ember 4, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Dinner, Dara L.

REGISTRATION NUMBER: 33,680

REFERENCE/DOCKET NUMBER: C70237

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5017

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 5150 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Ribes nigrum

STRAIN: Ben Alder

US-09-068-140A-14

```
Query Match      8.4%; Score 92; DB 4; Length 5150;
Best Local Similarity 49.3%; Pred. No. 3.4e-13;
Matches 347; Conservative 0; Mismatches 330; Indels 27; Gaps 3;

Qy 189 cgaggcgtgcggctgagccactgtcacacgacacatggtggcaggcgcggtggcagg 248
Db 3236 CGCGCGGTGCGATGATGCGCTTCACTTGGCAGTTCATGATGCGCGGTTCATCGCTGG 3295
Qy 249 gatcctgagcactgcgtgatgtacccatgcactgcctcaagaccgagcagtgatgta 308
Db 3296 ATCAATCGAGCATATGCGGATGATCCGGTGTGATACGCTTAAACTCCGATACAGGGTAT 3355
Qy 309 acagcctgaccagctgcccgcctatcgcaatggttggaggccctctggagattataag 368
Db 3356 TGGGTTCATGTTGGCT---CAATCCGCGGTCTCCGACAAGCCCTTGGGTGATACTGAA 3412
Qy 369 aacgagggcctatgagggcccatgagggcctgagggcctgaacgtcacagcaacaggcggcc 428
Db 3413 AGTTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGCGAGGACC 3472
Qy 429 tcccacgcccctttatttgcctgtctacgaaagttaaagacacatgagtgatgta 488
Db 3473 AGCTCAGGAGTGTATTTCTCGTTTACGATGTGTAAAGAGACTTTTCTCATGGTGA 3532
Qy 489 ccaccctgggggcaatagccatattgccaatggtgcggcgggtgtgtggcaacattact 548
Db 3533 TCCGAGCAATTCGGTGGCGACGCGTTCGGGGGTGTTCGGGACGGTGGCAA----- 3585
Qy 549 tcagtgcagccatgaaccctgcgaagtggtcaagcagagagatgcagatgtacaactc 608
Db 3586 --GCGACGCGGTGATACGCCGATGGATGTGTTGAACAGAGAGTTGCGAGTTGCAGAGCAG 3643
Qy 609 accataccacgggtgacagactgtgacggcagtggtgcaaaatgagggcggcgggc 668
Db 3644 TCCGTACAGGGTGTGTGTGATGCGTGGAGGAGGTGTGTGAGAGAGGATTTGGCGC 3703
Qy 669 cttttaccgcagctacacacccagctgacctgaacgttctcttccaaagccattcact 728
Db 3704 ATTTTACGCATCTTATCGAACAACCTGTGTCATGAATGCCCGGTTACGGCGGTTCACTT 3763
Qy 729 catgacctatgaatcctgcagagacactttaacccccagagacggttacaccccaagctc 788
Db 3764 CCCCACATATGAAGCCAGCAAGAAAGGTTGTGTGAGGTTGTGCGCGGAGACTTGCAGAACGA 3823
Qy 789 ccagctcctc-----ctggagcttgcagagagctgtagctgccgcagc 833
Db 3824 TGAGAAATTTGTAGTGCATGTACTGTGCTGGTCTGCTGCTGGAGCTTTGGCTGCAGTAGT 3883
Qy 834 cacaacccactggcgcttgcacacacactgctcacacccagg 877
Db 3884 AACCACTCCACTAGATGTGTCAAAACTCAGTTGTCAGTGGCCAG 3927

RESULT 2
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068.140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L. 33, 680
REGISTRATION NUMBER: C70237
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-9

Query Match      8.3%; Score 90.4; DB 4; Length 1311;
Best Local Similarity 49.1%; Pred. No. 5.7e-13;
Matches 346; Conservative 0; Mismatches 331; Indels 27; Gaps 3;

Qy 189 cgaggcgtgcggctgagccactgtcacacgacacatggtggcaggcgcggtggcagg 248
Db 72 CGCGCGGTGCGATGATGCGCTTCACTTGGCAGTTCATGATGCGCGGTTCATCGCTGG 131
Qy 249 gatcctgagcactgcgtgatgtacccatgcactgcctcaagaccgagcagtgatgta 308
Db 132 ATCAATCGAGCATATGCGGATGTATCCGGTGTGATACGCTTAAACTCCGATACAGGCTAT 191
Qy 309 acagcctgaccagctgcccgcctatcgcaatggttggaggccctctggagattataag 368
Db 192 TGGGTTCATGTTGGCT---CAATCCGCGGTCTCCGACAAGCCCTTGGGTGCTGATACTGA 248
Qy 369 aacgagggcctatgagggcccatgagggcctgaacgtcacagcaacaggcggcc 428
Db 249 AGTTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGCGAGGACC 308
Qy 429 tcccacgcccctttatttgcctgtctacgaaagttaaagacacatgagtgatgta 488
Db 309 AGCTCAGCAGTGTATTTCTCGTTTACGAGATGTGTAAAGAGACTTTTCTCATGGTGA 368
Qy 489 ccaccctgggggcaatagccatattgccaatggtgcggcgggtgtgtggcaacattact 548
Db 369 TCCGAGCAATTCGGTGGCGACGCGTTCGGGGGTGTTCGGGAGGTGGCAA----- 421
Qy 549 tcagtgcagccatgaaccctgcgaagtggtcaagcagagagatgcagatgtacaactc 608
Db 422 --GCGACGCGGTGATTTACCGGATGATGATGTTGTTGAGAGAGGATTTGAGAGAGGATTTGGCGC 479
Qy 609 accataccacgggtgacagactgtgacggcagtggtggcaaatgagggcggcgggc 668
Db 480 TCCGTACAAAGGTTGTGTGATTTGCGTGAGGAGGTTGTGTGAGAGAGGATTTGGCGC 539
Qy 669 cttttaccgcagctacacacccagctgacctgaacgttctcttccaaagccattcact 728
Db 540 ATTTTACGCATCTTATCGAACAACCTGTGTCATGAATGCCCGGTTTACGGCGGTTCACTT 599
Qy 729 catgacctatgaatcctgcagagacactttaacccccagagacggttacaccccaagctc 788
```



```

; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-028-155-16

```

[illegible]

```

RESULT 7
US-09-165-264-7
; Sequence 7, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagancorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer sequence
US-09-165-264-7

```

[illegible]

```

RESULT      8
US-09-165-264-13
; Sequence 13, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 13
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence
US-09-165-264-13

```

[illegible]

```

RESULT          9
US-09-165-264-11
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiyiah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165.264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer sequence
US-09-165-264-11

```

[illegible]


```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..833
; US-08-510-646B-3

Query Match 4.6%; Score 50.6; DB 3; Length 833;
Best Local Similarity 53.0%; Pred. No. 0.0018;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

QY 3 ggagttggagggcggtgctgctgagcggtggtgagggggcgccgagggcgcccgggcg 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 GCGTTCGCGCAGGGTGGTGGCGCGGTGCGGTCCGGCGCGCGCGCGCGCGCGCG 728

QY 63 gagcc-----ccggggagtcgagcggtgctgagcggtggtgagcgggcgcccg 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 727 CGGCCAGGACCGGGTCTCGAGCGCGGTGAGCGCCACCGTGAAGCGTCCGCGGGGCTGT 668

QY 119 gggcgcgcgcggggagggcgccggtgagcgcccggtgagcgcccggtacgacagatccggactccg 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CGCGGGCAGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 608

QY 179 gcccgactacgagcggtgctgagcggtgagcgcccggtgagcgcccggtacacacgacatggtggcagggcg 238
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 GCTGCTTCCAGCGCGCGTCCGCGCGCGTCCGCGCGCGTCCGCGCGCGCGCGCG 548

QY 239 ccgtggcaggg 249
| ||| |||
Db 547 CGAGGCGGGG 537

RESULT 12
US-09-231-818-3/c
; Sequence 3, Application US/09231818
; Patent No. 6171846
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231,818
; FILING DATE:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/403,852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..833
; US-09-231-818-3

Query Match 4.6%; Score 50.6; DB 4; Length 833;
Best Local Similarity 53.0%; Pred. No. 0.0018;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

QY 3 ggagttggagggcggtgctgctgagcggtggtgagggggcgccgagggcgcccgggcg 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 GCGTTCGCGCAGGGTGGTGGCGCGGTGCGGTCCGGCGCGCGCGCGCGCGCGCG 728

QY 63 gagcc-----ccggggagtcgagcggtgctgagcggtggtgagcgggcgcccg 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 727 CGGCCAGGACCGGGTCTCGAGCGCGGTGAGCGCCACCGTGAAGCGTCCGCGGGGCTGT 668

QY 119 gggcgcgcgcggggagggcgccggtgagcgcccggtgagcgcccggtacgacagatccggactccg 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CGCGGGCAGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 608

QY 179 gcccgactacgagcggtgctgagcggtgagcgcccggtgagcgcccggtacacacgacatggtggcagggcg 238
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 GCTGCTTCCAGCGCGCGTCCGCGCGCGTCCGCGCGCGTCCGCGCGCGCGCGCG 548

QY 239 ccgtggcaggg 249
| ||| |||
Db 547 CGAGGCGGGG 537

RESULT 13
US-08-403-852D-1/c
; Sequence 1, Application US/08403852D
; Patent No. 5891695
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2002, 00:28:35 ; Search time 1719.3 Seconds

(without alignments)

8596.037 Million cell updates/sec

Title: US-09-870-113-1

Perfect score: 1095

Sequence: 1 atgagttggaggcgagg.....agtggaggcgcaagtga 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	692.2	63.2	888	9	AL530804
2	637.4	58.2	747	10	BI827322
3	615.8	56.2	1131	10	BG295496
4	601.4	54.9	681	9	AV704087
5	587	53.6	826	10	BI254253
6	586	53.5	986	10	BG331197
7	560.4	51.2	645	10	BG964218
8	543	49.6	855	10	BI412175
9	534.6	48.8	817	10	BI103329
10	532.2	48.6	576	10	BI057218
11	517	47.2	927	10	BG972351
12	507.6	46.4	574	10	BI118580
13	507.4	46.3	908	10	BG697119
14	502.4	45.9	560	10	BI794616
15	502.2	45.9	598	10	BE913718
16	473	43.2	808	10	BG918082
17	472.6	43.2	531	9	AW915209

18	470.6	43.0	563	10	BF193023
19	468	42.7	581	9	AW341177
20	453.6	41.4	830	10	BF143133
21	453.4	41.4	578	10	BM483098
22	441.4	40.3	480	10	BI045863
23	436	39.8	1295	11	AK015770
24	435	39.7	451	10	BF515180
25	433.8	39.6	529	9	AA431276
26	433	39.5	501	9	AW211366
27	432	39.5	527	9	AI797090
28	432	39.5	526	9	AI797100
29	430.2	39.3	1209	11	AK006155
30	429.6	39.2	741	9	AL563845
31	427.2	39.0	709	9	AJ397503
32	419.4	38.3	573	10	BI541703
33	417.8	38.2	483	9	AA061624
34	409.8	37.4	697	10	BI854643
35	405.6	37.0	563	10	BF078769
36	398.4	36.4	541	9	AA234031
37	398.2	36.4	967	10	BI739710
38	393.6	35.9	453	10	BF916224
39	391	35.7	488	9	AI797101
40	387	35.3	718	10	BJ041453
41	383.2	35.0	423	10	BI339456
42	382.8	35.0	425	10	BI304641
43	381.6	34.8	848	9	AI743110
44	371	33.9	434	10	W53634
45	368.2	33.6	491	10	BF076673

ALIGNMENTS

RESULT 1

AL530804

LOCUS

DEFINITION

AL530804 LTI_NFL001_NBC4 Homo sapiens

prime, mRNA sequence.

AL530804

VERSION

AL530804.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 888)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="LTI_NFL001_NBC4"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT

168 a 244 c 322 g 139 t 15 others

888 bp mRNA linear EST 13-FEB-2001

AL530804 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YH17 5

prime, mRNA sequence.

AL530804

VERSION

AL530804.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 888)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="LTI_NFL001_NBC4"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

ORIGIN

Query Match 63.2%; Score 692.2; DB 9; Length 888;
Best Local Similarity 97.1%; Pred. No. 6.1e-128;
Matches 748; Conservative 8; Mismatches 8; Indels 6

		Matches	748; Conservative	8; Mismatches	Indels	6; Gaps
QY	1	atgagtgaggagggcggggtcgtgcgcgtgtgcgcggggggcgccgagcgccgagggcccggg	60			
Db	122	ATGAGTATGAGGAGGGCGGGGTGCTGCGGTGTGCGCGGGGGCGCGGCGCAGGCGCCCGG	181			
QY	61	cggagccccgggagtgctggcgctgcctggaacgggtggctgcagcgggcgctggggccgggg	120			
Db	182	CGGAG--CCGGGAGTCCGGCTGCTGGACGGTGGCTGCAGCGGGGCGCTGGGCGCGGG	239			
QY	121	gccggcggggagagccgggacctcaggcccccggtagcagcaaatcgggactccggc	180			
Db	240	GCCGGCGCGGGGAGGCGGGGGCTGCAG-CCCCGGTAGCAAGATCCGGACTCCGG	298			
QY	181	ccggactcagggcgctgcggg-ctggagccactgtcacccagcatggtggcagcgcg	239			
Db	299	CCGGACTACGAGCGCTGCCGNGCTGGAGCCACTGTCCACACGCACATGTTGTCAGGCGC	358			
QY	240	cgtaggcaggaatcctggagcaactgcgtgatgtaccoccatcgactgcgtcaagaccggat	299			
Db	359	CGTGGCAGGATCTCGGAGCATTGCGTNAWGTATCCCATCGACTGCTCAAGACCCGGAT	418			
QY	300	gcagagctctacagcctgaccagcgtcccgctatcgcaatggttgagagccctctgag	359			
Db	419	GCAGAGTCTACAGCTTGACCCAGCTGCCCGCTATCGCAATGTTGGAGGCCCTCTGGAG	478			
QY	360	gattataagaacggagggcctatggaggcccatggaggggctgaacgctcacagcaacagg	419			
Db	479	GATTATAAGAACAGGGGCTATGGAGGCCCATGGAGGCCATGAGGGGCTGACAGCTCACAGCAACAG	538			
QY	420	cgagggcctgccacgcgccttatttgcctgcctacgaaaagttaaaaaagacattgag	479			
Db	539	CGCAGGGCCTGCCACGCGCCTTATTTTGTGCTGCACGAAAAGTTAAAAAGACATGTAG	598			
QY	480	tgatgtaatccaccctggggcgaatgacatattgccaatggtgcgcggcggtgtgtgc	539			
Db	599	TGATGTATCCACCTCGGGGGCAATAGCCATATGCAATGGTTCGGCGCGGGTGTGTGGC	658			
QY	540	aacattactctatgatcgagccatgaacctgcggaagtggtaacgacagagtagtcagat	599			
Db	659	AACATTACTTCATGATGCAGCCATGAACCTCGGGAAGTGTCAAGSAGAGGATGCAGAT	718			
QY	600	gtacaactcacatcacacgggtgacagactgtgtacggggcagttt-ggcataaatgaag	658			
Db	719	GTACAACCTCACATACCACCGGTGCACACTGTGTACGGGCGAGTGTGGGCAAAATGAAG	778			
QY	659	ggcgccggggcctttaccgcgagctacacacccagctgacctgaacgttcccttccaag	718			
Db	779	GGCGCGGGGCTTTTACCGCAGCTACACACCCAAATGATKACGTTCTCTTCCAG	838			
QY	719	ccattcaccttca-tgacctatgaattcttcgaggagcacttttaacccccca	767			
Db	839	CCATTTACTTCAATGACCTATGAATTCCTCGAGAGCCCTTTTAACCCCCA	888			

RESULT 2

BI827322	747 bp	mRNA	linear	EST 04-OCT-2001
LOCUS	BI827322	603077989F1	NIH_MGC_119	Homo sapiens CDNA clone IMAGE:5169694 5',
DEFINITION	mRNA sequence.			

ACCESSION	BI827322
VERSION	BI827322.1
KEYWORDS	EST.
	GI:15938872

SOURCE human.

ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 747)
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 747)

[illegible]

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
plate: LLaM11422 row: a column: 23
High quality sequence stop: 743.

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Charitable Contributions</i>	16. <i>Charitable Contributions</i>
17. <i>Volunteer Work</i>	17. <i>Volunteer Work</i>
18. <i>Substance Use</i>	18. <i>Substance Use</i>
19. <i>Mental Health</i>	19. <i>Mental Health</i>
20. <i>Life Satisfaction</i>	20. <i>Life Satisfaction</i>

```
1. 747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
note="Organ: brain; Vec
```

/note=Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned [EcoRV site is destroyed upon
 cloning]. Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note
 this is a NIH MGC Library."

BASE COUNT	188 a	206 c	201 g	151 t	1 others
ORIGIN					

Query Match	58.2%;	Score 637.4;	DB 10;	Length 747;
Best Local Similarity	98.1%;	Pred. No. 4.8e-117;		
Matches 666:	Conservative 0;	Mismatches 11;	Indels 2;	

Qy	275	ccatcgactgctcaagagaccgcatgacagtgctacagcctgaccagctgcccgtatc	334
Db	61	CTACCGGCTGATTTTCCCCGGATGCAGAGTCTACAGCCTGACCCAGTCCCGGTATC	120
Qy	335	gcaatgtgttgaggccctctggaggattataaagaacgaggccctatggagcccattga	394
Db	121	GCAATGTGTTGGAGGCCCTCTGAGGAGTATTAAAGACGGAGGGCCCTATGGAGGCCCATGA	180
Qy	395	gggggctgaacgtcacagcaacaggcgagggcctgccacgcccctttatttggctgct	454
Db	181	GGGGCTGAACGTCACAGCAACAGGCGCAGGGCTGCCACAGCCCTTATTTTGCGCTGCT	240
Qy	455	acgaagaattaaaaagacattgagtgtgtaataccacctgggggcaatagccatttg	514
Db	241	ACGAAAAGTTAAAAAAGACATTGAGTGTGATTAATCCACCCTGGGGCAATAGCCATATTG	300
Qy	515	ccaatgctgcgcgggtgtgagcaacttctcatgatcagccatgaaccctgcgg	574
Db	301	CCAATGCTGGCGGGGTGTGGCAACATTACTTTCATGATGCAGGCATGAACCTTGGG	360
Qy	575	aagtgtgccaagcagaggtgcagatgtcaaacctcacataccacccgggtgacagactgtg	634
Db	361	AAGTGGTCAAGCAGAGGATGCAGATGTACAACCTACCATACCACCGGTGACAGACTGTG	420
Qy	635	tacgggcagtgctggaaaaatgaaggggcggggcctttaccgcagctacacaccacgc	694
Db	421	TACGGGCAGTGTGGCAAAATGAAGAGGGCGGGGCCCTTTTACCGCAGCTACACCACCCAG	480
Qy	695	tgaccatgaacgttctcttccaaagccattcaactcatgacctatgaattctctcaggagc	754
Db	481	TGACCATGAACGTTCTTTTCCAAGCCATTCACTTCATGACCTATGAATTCCTGCAGGAGC	540
Qy	755	actttaacccccagagacggtacaaacccaagctcccaacgctccctctctgaggttgcgac	814

This clone is available at CHGC in Shanghai.

BI254253	826 bp	mRNA	linear	EST 17-JUL-2001		
LOCUS	602974729F1	NIH_MGC_12	Homo sapiens	CDNA clone IMAGE:5114190 5',		
DEFINITION	mRNA sequence.					
ACCESSION	BI254253					
VERSION	BI254253.1	GI:14806485				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 826)					

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia, Eutheria, Artiodactyla, Primates, Carnivora, Hominoidea.
1 (bases 1 to 826)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Prepared by: Incyte Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LLNL at:

Plate: LLAM11277 row: i column: 07

FEATURES

source
I. .826
/organism="Homo sapiens"

/db_xref="taxon:9606"
/db_xref="taxon:9606"

```
/clone_lib="NIH_MGC_12"
```

```
/lab_host="DH10B"
```

```

/site 2: salI: cloned unidirectionally. Primer: oligo
/note= origin: celvix; vector: pcmv-sfork10; site_1: no

```

Average insert size 1.4 kb. Library prepared by Life Technologies[®]

BASE COUNT	211 a	233 c	208 g	174 t

Query Match 53.68; Score 587; DB: 10; Length 826;

Best Local Similarity 99.8%; Pred. No. 5.2e-10;
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps

1000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

QY 55/ cagccatgaaccctgcggaagtgggtcaagcagaggatgcagatgtaccaactcaccataacc 616

Db 60 CAGCCATGAACCCCTGCGGAAGTGGTCAAGCAGAGGATGCAGATGTACAACCTCACCATACC 119

Qy 617 accgggtgacagactgtgtacgggcagtgtggcaaaatgaaggggcccgtttacc 676

Db 120 ACCGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAAGGGCCGGGGCCTTTACC 179

[illegible]

.....

[illegible]

DB 240 ATGAATTCCTGCAGGAGCACTTTAAACCCCGAGAGACGGTACAACCCACAGCTCCACGTTCC 299

```
QY 797 tctctgagcttgcgagagctgtagctgcgcagccacacccccactggacgtttgca 856
|||||
Db 300 TCTCTGGAGCTTGGCAGGAGCTGTAGCTGCCGACGCCACACCCACTGGAGTTGCA 359
QY 857 aaacactgtcacaacccccaggagctccttgcttggaactcacacattacaggacatatca 916
|||||
Db 360 AAACACTGCTCAACACCCAGGAGTCTTGCTTTGAATCACAATACAGGACATATCA 419
QY 917 caggcatgctagcttcagacacggttatataagtaggtggggtgaccgcctatttcc 976
|||||
Db 420 CAGGCATGGCTAGTGCTTTCAGGACGGTATATCAATAGGTGGGTGACCGCTATTTCC 479
QY 977 gaggggtgcagccagagtaattaccagatccctccacagccatcgatgctctgtg 1036
|||||
Db 480 GAGGGGTGCAGGCCAGAGTAATTTACCAGATCCCTCCACAGGCATCGCATGTCTGTGT 539
QY 1037 atgagttcttcaataacacctaatacctaagaagcaagaagtgagggtggcgaagtga 1095
|||||
Db 540 ATGAGTTCTTCAATAACCTAATCACTAAAAGGCAAGAAGAGTGAGGGCTGGCAAGTGA 598

RESULT 6
BG331197 986 bp mRNA linear EST 27-FEB-2001
LOCUS 602431907F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4549801 5',
DEFINITION mRNA sequence.
ACCESSION BG331197
VERSION BG331197.1 GI:13137635
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 986)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCRD/DPF/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1241 row: m column: 02
High quality sequence stop: 587.
High quality sequence start: 587.
Location/Qualifiers
1. 986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4549801"
/tissue_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 268 a 277 c 248 g 193 t
ORIGIN
```

Query Match 53.5%; Score 586; DB 10; Length 986;
Best Local Similarity 97.5%; Pred. No. 8.5e-107;
Matches 595; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
QY 484 gtaatcacacctgggggcaataagccattatgccaatgggtgcgcgggtgtgtggaaca 543
|||||
Db 2 GTAATCCACCTGGGGCAATAGCCATATTGCCAATGGTGGCGCGGGGTGTGGCAACA 61
QY 544 ttacttcagtgcagcccatgaacctgggaagtggtcaagcagaggaatgcagatgtac 603
|||||
Db 62 TTACTTTCATGATGCAGCCATGAACCCCTGCGAAGTGGTCAAGCAGAGGATGCAGATGTAC 121
QY 604 aactcaaccatacacacggggtgacagactgtgtacgggagctgtggcaaaatgaaggggcc 663
|||||
Db 122 AACTCACCATTACCACCGGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAAGGGGCC 181
QY 664 ggggacctttaccgcagctacacccacccagctgacatgaacgttctcttcccaagccatt 723
|||||
Db 182 GGGGCCCTTTACCGCAGCTACACCCAGCTGACCATGNACGTTCTTTTCCAAAGCCATT 241
QY 724 cacttcagctatgaattcctgcaggagcactttaacccccacagagcaggtacaaccca 783
|||||
Db 242 CACTTCATGACCTATGAATTCCTGCAGAGGACACTTTAAACCCCCAGAGCGGTACAACCCA 301
QY 784 agctccacagctcctctctgagcttgcgaggagctgtagctgccgacgacacacccca 843
|||||
Db 302 AGCTCCACGCTCCTCTGGAGCTTGCAGAGGAGCTGTAGCCTGCCGACGCCACAACCCA 361
QY 844 ctggagctttgcaaaacactgctcaacacccagagagctccttggcttgaactcacacatt 903
|||||
Db 362 CTGGACGTTTGGAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAATCACAAT 421
QY 904 acaggacatatcacaggcagctgctagctgccttcagacgcttatcaagttagtggtgtg 963
|||||
Db 422 ACAGGACATATCACAGGCATGGCTAGTGCCTTCAGGACGGTATATCAAGTAGTGGGTG 481
QY 964 accgcctattccaggggtgcagccagagtaattaccagatccctccacagccatc 1023
|||||
Db 482 ACCGCTATTTCCGACGGGTGCAGGCCAGAGTAATTTACCAGATCCCTCCACAGCCATC 541
QY 1024 gcatgtctgtatgagttcttcaaatcctaataacacaaagcgaagaagtagtgag 1083
|||||
Db 542 GCATGCTGTGTATGATGCTTCAANTACCTAATACCTAAAGGCAAGAACACTGGAGG 601
QY 1084 gctggcgaagt 1093
|||||
Db 602 GCTGGCAAGT 611

RESULT 7
BG964218 645 bp mRNA linear EST 12-JUN-2001
LOCUS 602828996F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983815 5',
DEFINITION mRNA sequence.
ACCESSION BG964218
VERSION BG964218.1 GI:14351855
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 645)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10988 row: p column: 24
High quality sequence start: 2
High quality sequence stop: 632.
```

FEATURES source Location/Qualifiers
1..645
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4983815"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DHI0B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 147 a 189 c 193 g 116 t
ORIGIN

Query Match 51.2%; Score 560.4; DB 10; Length 645;
Best Local Similarity 92.1%; Pred. No. 9.9e-102;
Matches 591; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 129 cggggagcgggctcaggccccgggtacgacaagatccggactccggccggacta 188
Db 4 CGCGTACGGGGGGCTATPACAGCCCCCTGTACGGCTGGATCGGAGTCGGGCCCGGAATA 63

QY 189 cgaggcgctgcggctggagccactgtaccacgcacatggtggtggcggcagg 248
Db 64 CGAAGCGCTGCCGGTGGAGCCACTGTACCACGCACATGTTGGGGGGCGCGTGGCAGG 123

QY 249 gatcctggagcactggtgatatacccatcactgcgtcgaagaccgagatcacagtct 308
Db 124 GATCTGGAGCATYTCGGTGTATGCCGATCGACTGCGTCAAGACCCGGATCGAGAGCT 183

QY 309 acagcctgaccagctgcccgtatcgcaatgtgttgaggccctctggaggattataag 368
Db 184 ACAGCGCTGACCCAGCGCCGCTATCGGAACCTGTTGGAGGCTCTCTGGAGAATCATGAG 243

QY 369 aacgagggcctataggcccatgaggggctgaacgtcaacagcaacagggcgagggcc 428
Db 244 GACAGAGGCCCTGTGGAGGCCCATGCGGGGGCTGAACGTCACAGCAACAGCGCGGGGCC 303

QY 429 tgccacgcctttatttgcctgctacgaaaagttaaagaagacatigagtgtgtaat 488
Db 304 TGCCACAGCCCTTATTTTGGCTGTACGAAAAGTTAAAAAGACATTTGAGTACGTAAAT 363

QY 489 ccaccttgggggcaatagccatattgccaatggtggtggccgggtgtgtggcaacattact 548
Db 364 CCACCCAGGGGCAATAGCCATATTGCCAATGTTGCCAATGTTGCCAGCGGATGTGTGGCACAATTACT 423

QY 549 tcatatcagccatgaaccctcggaagtgggtcaagcagagggatgcagatgtacaactc 608
Db 424 TCATGATCAGCCATGAATCCAGCGGAAGTCGTCAAGCAGAGGATGCGAGATGTACAACCTC 483

QY 609 accataccacgggtgacagactgtgtacggcgaagtgtggcaaatgaaggggcgagggc 668
Db 484 GCCGTACCCACCGGTGACAGACTGTGTTGGGCGAGTGTGGCAAAATGAAGGGGGCGGGGC 543

QY 669 cttttaccgcagctacacaccacagctgaccatgaacgttctcttcccagccattcactt 728
Db 544 CTTTATCCGCGAGCTACACGACCCAGCTGACCATGAATGTCTCCCTTCCCAAGCCATTCATT 603

QY 729 catgacctatgaattcctgcaggagcacttttaacccccagag 770
Db 604 CATGACCTATGAGTTCTCTGCAAGAGCACTTTAAAGCCCCCAGAG 645

RESULT 8
BI412175
LOCUS BI412175 855 bp mRNA linear EST 14-AUG-2001
DEFINITION 602965377F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5120755 5',
mRNA sequence.
ACCESSION BI412175
VERSION BI412175.1 GI:15173098
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 855)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
CDNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM11294 row: j column: 20
High quality sequence start: 7
High quality sequence stop: 809.
Location/Qualifiers
1..855
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5120755"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DHI0B (phage-resistant)"
/note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCCTCTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 209 a 233 c 241 g 172 t
ORIGIN

Query Match 49.6%; Score 543; DB 10; Length 855;
Best Local Similarity 88.2%; Pred. No. 3e-98;
Matches 738; Conservative 0; Mismatches 85; Indels 14; Gaps 13;

QY 229 gtggcaggcgccgtggcaggatcctggagcactgcgtgtatcccccatcgactgcgtc 288
Db 18 GTGGCGGGCGCGTGGCA-GGATCCTGGAGCATTCGCTGATGCCCATCGGGTGGCTC 76

QY 289 aagaccggatgcagagctcacagcctgaccag-ctgcccgctatcgcaatgtgttga 347
Db 77 AAGACCCGGATGACAGACCTACAGCCTGACCCAGCCCGCTATCGGAACGTTGTGGA 136

QY 348 ggcctctggaggattataaagcggagcctatgagggcccatgagg-gctgaacg 406
Db 137 GGCTCTCTGGAGATCATGAGACAGAGGGGCTGTGGAGGCCCATGCGGGGTGCTGAACG 196

QY 407 tcacagcaacaggcgagggcc-tgccacagccct-ttatcttgcctgctacgaaaagt 464
Db 197 TCACAGCAACAGCGCGGGGCGGTGCCACGCCCTCGTATTTTTCCTGCTACGAAAGTT 256

QY 465 aaaaaagaca-ttgagtgtatgaatccacccctgggggccaatagccatattgccaatggt 523
Db 257 AAAAAAGACAGTTGAGTGACGTAATCCACCAGGGGCAATAGCCATATTCCTAATGGTG 316

QY 524 cgccgggtgtgagcaacttacttatcatgacccatgagccatgaacccctgcgaagt-ggtc 582
Db 317 CAGCCGATGTGGCGACATATTCTTATGATGAGCCATGAATCCAGGAAAGTCGTTTC 376

QY 583 aagcagaggatgcagatgtacaactcaccatccacccgggtgacagactgtgtacgggca 642

```

|||||
377 AACGAGAGGATGAGATGTAACTCCGCTACCACCGCTGACAGACTGTGTTCGGCA 436
QY 643 gtgtgcaaatgaaggccgggctttttaccagctacac-caccagctgacct 701
Db 437 GTGTGGCAAAATGAAGGGCCGGGGCTTTTACCGCAGCTACACTGACCCAGCTGACCAT 496
QY 702 gaacgttccttccaagccattcacttcacatgacctatgaattcctcagagcaccttaa 761
Db 497 GAATGTCCCTTCCAAGCCATTCACTTCATGACCTATGAGTTCTTGGCAAGACACTTTAA 556
QY 762 ccccccagagcgtgtacaaacccaagctcccccactcctctctgtgagcttgcgcagagctgt 821
Db 557 ACCCCAGAGACGGGTACAAGGCCAGCTCCCATGTCTGTGTGGGCCCTG-GCAGGAGCAGT 615
QY 822 agtcgcagcagcacacccccactgagcgtttgcacaaacactgctcaaacccagg--ag 879
Db 616 AGCTGCCGCCGCCACAAAGCCCACTGGACGTTTGTAAACACACTGTCTCAACAGCCAGGAAT 675
QY 880 tccttgctttgaactcacacattacagggacatatcaca-ggcatggctagtgccttca- 937
Db 676 CCCTGGCTTTGAACACTCAAACTTACTGGACACATCACAGGGCATTTGCTAGTGCCCTTCAG 735
QY 938 ggaacgttatcaagtagtggtggggtgaccgcctatttccagggggtgcagggccagagtaa 997
Db 736 GGACGGGTCTATCAAGTAGGGGGTGTGACTGCTTACTTCCAGGGGGTGCAGGGCAGAGTCA 795
QY 998 ttaccagatcccctcc-acagccatcgcatgctgtgtatgagtcttccaatatc 1053
Db 796 TTTACCAGATCCCTCCAAACAGCATCGCATGGTCTTGTGTATGAATTTCAAANATC 852

RESULT 9
BI103329
LOCUS 817 bp mRNA linear EST 26-JUN-2001
DEFINITION 60288838F1 NCI_CGAP_Kid14 Mus musculus cdna clone IMAGE:5043992
5', mRNA sequence.
ACCESSION BI103329
VERSION 1 GI:14554222
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 817)
COMMENT NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1120 row: 1 column: 09
High quality sequence stop: 777.
FEATURES
Location/Qualifiers
1..817
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5043992"
/lab_host="NCI_CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. ]"
147 a 222 c 316 g 132 t
BASE COUNT
ORIGIN

```

```

Query Match 48.8%; Score 534.6; DB 10; Length 817;
Best Local Similarity 90.7%; Pred. No. 1.4e-96;
Matches 603; Conservative 0; Mismatches 59; Indels 3; Gaps 3;

QY 29 gtgtggcgggggcccggcagggcccgggcgagccccggggagccccggggagctcggcgctgtg 88
Db 153 GTGTGGGAGTGAACCATGCTGCTGGGCCCGGGGAGAGCCCGGGGAGTTCGGCGCTGTGG 212
QY 89 acgggttgctgcagcggggcggtggcgccggcgccggcgccggcgccggcgccggcgccgctgca 148
Db 213 ACGGTGGCTGTCAGCGGGGGGTGGCGCGGGGGCGCGCGGGAGAGCGGGGGCTTATC 272
QY 149 gcccccggttacacagaatcccgactcccgccggcgccggcgactacgagggcctcccgctgag 208
Db 273 AGCCCCCTGATCGGCTGGATCCGGAGTCCCGCCCGCGAATAGGAAGCGCTGCCGGCTGGAG 332
QY 209 ccactgtcaaccacgcacatggtggcaggcccggtggcaggatcctggagcactgcgtga 268
Db 333 CCACCTGTCAACCGACACATGCTGGCGGCGCGCTGGCAGGAGTCTCTGGAGCATTTGCGTGA 392
QY 269 tgtaccaccatcgactgctcaagaccggatgcagagctacacgctgagccagctgccc 328
Db 393 TGTACCCGATCGACTGCGTCAAGACCCGATGCAGAGCCCTACAGCCCTGACCCAGCCGCC 452
QY 329 gctatcgcaatgtgttgaggccctctggaggattataagaacgagggcctatggaggc 388
Db 453 GCTATCGGAACGTGTGGAGGCTCTCTGGAGATCATGAGGACAGAGGGGCTGTGGAGGC 512
QY 389 ccatgagggggctgaacgtcacagcaacagggcgaggcctgcccagcgccttattttg 448
Db 513 CCATGCGGGGCTGAACGTCAACAGACAGCGCGGGGCGCTGCCCGCCCTCTATTATTG 572
QY 449 cctgctacgaaaagttaaaaaaacattgagtgatgtaataccacctgggggcaatagcc 508
Db 573 CCTGCTACGAAAAGTTAAAAAGACATTTAGTGACGTAAATCCACCCAGGGGCAATAGCC 632
QY 509 atattgccaatgg-tgcgcggggtgtgtgcaacattac-ttcagtatgcagccatgaa 566
Db 633 ATATTGCCAATGTTGCGCGCGGATGTGGCGACATTTCTTCTATGATGCGAGCATGAA 692
QY 567 cctctgcggaagtgtcaagcagagagtgatgtagatgtacaactcaccatccaccgggtgac 626
Db 693 TCCAGCGGAAGTGTCTCAAGCAGAGGATGCAGATGTACAACTCGCGCTACCAACCGCGTAC 752
QY 627 agactgtgtacggcgagtg-tggcaaaatgaaggcgccgggctttttacgcagctaca 685
Db 753 AGACTGTGTTCGGGCGAGTGTGGCAAAATGACAGGGGGGGGCTTTTAAACCGCAGCTACA 812
QY 686 ccacc 690
Db 813 CGAAC 817

RESULT 10
LOCUS BI057218/c
DEFINITION QV3-CN0338-230201-558-a01 GN0338 Homo sapiens cdna, mRNA sequence.
ACCESSION BI057218
VERSION BI057218.1 GI:14464748
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```



```
Db 250 GCAAGAGCACTTTAAACCCAGAGACAGGTACAACCCAGCTCCCATGTGCTCTGTGGGC 309
QY ttgcagagagctagctccgagacacacacccactgagctttgcaaaacactgct 866
Db 310 CTGCGCAGGAGCAGTAGCTCGCGCCGACACACCCCACTGGACGTTTGTAAACACACTGCT 369
QY caacacccaggagtccttggctttgaactcacacattacagagacatcacagggcatggc 926
Db 370 CAACACCCAGGAATCCCTGGCTTTGAACCTCAAAACATTACTGGACACATCACAGGCATGGC 429
QY tagtgccttcaggagcgatatacaagtagtgggtgggtgacccgctatttccagaggggtgca 986
Db 430 TAGTGCCTTCAGGACGCTCTATCAAGTAGCGGGGTGACTCTTACTTCCGAGGGGTGCA 489
QY ggcagagtagtaattaccagatccctccacagccatcgatggtctgtgtatgattctt 1046
Db 490 GGCAGAGTCAATTTACCAAGATCCCTCCACAGCCATCGATGGTCTGTGTAAGAATTTT 549
QY 1047 caatacctaatacctaaaaggcaagagagtgagggctggcgaagtga 1095
Db 550 CAATACCTAATCAAAAGCGGCAAGAGGAATGGAGGCGAGGCAAGTGA 598

RESULT 12
LOCUS BI185580
DEFINITION UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
ACCESSION UNL-P-FN-cy-f-02-0-UNL 3', mRNA sequence.
VERSION BI185580
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
TITLE 1 (bases 1 to 574)
JOURNAL Caetano,A.R., Johnson,R.K. and Pomp,D.
COMMENT Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA-No.

FEATURES
source Location/Qualifiers
1..574
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-cy-f-02-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DHI0B (Life Technologies)"
/notes="Vector: pT7M3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996.
TAG_SEQ=None found"
```

```
BASE COUNT 132 a 171 c 165 g 104 t 2 others
ORIGIN
Query Match 46.4%; Score 507.6; DB 10; Length 574;
Best Local Similarity 95.3%; Pred. No. 3.1e-91;
Matches 544; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
QY 169 ccgaactccgcgcgagactacagagcgctgcgg-ctgagagccactgtcacacgacacat 227
Db 5 CCGCTCFCGCGCCAGAGCTACGAGGCGCTGCCGNGCTGGAGCCACTGTGCACAGCACAT 64
QY 228 ggtggcagggccgctggcagggatccctgagcactgcgtgatgtaccccatcgactgcgt 287
Db 65 GGTGGCGGCGCCGTGGCAGGAATCCCTGGAGCACTGCGTGATGTACCCCATCGACTGGGT 124
QY 288 caagaccggagtgagagtgctacagcctgagcccgctgcccgcgtatcgcaatgtgtgga 347
Db 125 CAAGACCCGGATGCAGAGCCCTGCAGCCCTGATCCCGCCGCCGCTATCGCAATGTGTTGGA 184
QY 348 ggcctctgagagattataagaacgagggcgctatgaagcccatgagggcgctgaacgt 407
Db 185 GGCCTCTGGAGGATTATAAGACGGAGGCGCTGTGGAGGCCCATGCCGGGGGCTGAACGT 244
QY 408 cacagcaacgagcgagggcgctgccacgcccctttatttgcctgtctacgaaaagttaa 467
Db 245 CACAGCAACAGCGCAGGGCCTGCCAGCCCTCTATTTTGCCTGTATGAAAAGTTAAA 304
QY 468 aaagacattgagtgatgtaattccaccctgggggcaatagccatattgccaatgtgoggc 527
Db 305 AAAGACATTGAGTGATGTAATCCACCCTGGGGGCAATAGCCATATTGCAATGTGTGGCGC 364
QY 528 cgggtgtggcaacttacttcagtgcagccataaacctcggaagtggtaagca 587
Db 365 CGGTGGGT-ANAACTTACTTCAATGATGAGCCATGATCCAGAGAGTGGTCAAGCA 423
QY 588 gagatgcagatgtacaactcaccatcacccgggtgacagactgtgtacggcagtggtg 647
Db 424 GAGGATGCAGATGTACAACCTCACCCCTACCCGGGTGACAGACTGTGTACGGCAGTGTG 483
QY 648 gcaaatgaagggcgccgggctttaccgcagctacacacccagctgacacatgaacgt 707
Db 484 GCAAAACGAGGGCGCGGGCGCTTTTACCGCAGCTACACCACCAGCTCACCATGAACGT 543
QY 708 tcctttcaagcattcattcatgacctat 738
Db 544 TCCCTTCCAAAGCCATCCACTTCAATGACCTAT 574

RESULT 13
LOCUS BG697119
DEFINITION 602660340F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803681 5',
mRNA sequence.
ACCESSION BG697119
VERSION BG697119.1 GI:13962975
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 908)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 560)
AUTHORS	Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Secear, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blisstein, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.
TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished (2000)
COMMENT	Other_ESTs: ic92e02.x1 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium

Contact: Douglas Melton, Klaus H. Raescher, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 430.
 Location/Qualifiers
 1..560

QY	215	tcaccacgcacatggtggcaggcccgctgacagatcctggagcactgcgtgatacc	274
Db	1	TCACACGCACATGGTGGCGGCGCGTGGAGCATCTGGAGCATTCGGTGTATGACC	60
QY	275	ccatcgaactgcgtcaagaccgcgagatgcagaagtctaacgcctgaccacagctgcccgctatc	334
Db	61	CGATCGACTGCGTCAAGACCCGATCGAGACCTACAGCGTGACCCAGCGCCCGGTATC	120
QY	335	gcaatgtttggaggccctctgagatataaagaacgagggcctcatgaggcccatga	394

```
Db 121 GGAACCTGTTGGAGGCTCTCTCGAGAATCATGAGGACAGAGGCGCTGTGGAGGCCATGC 180
QY 395 ggggctaaactcagcaacagcgcaagcgctgcccacccctttatttgcctgt 454
Db 181 GGGGCTGAACGTCAAGTAACAGGCGCGGGCGCTGCCACGCCCTCTATTTGGCTGCT 240
QY 455 acgaaaagttaaaaaagacattgagtgatgtaattccacccctgggggcaatagccatttg 514
Db 241 ACGAAAAGTTAAAAAGACATTGAGTGAGCTAATCACCCAGGGGCAATAGCCATATTG 300
QY 515 ccaattggtggcgccgggtgtgtggaacattacttacttcagtcagccatgaacccctgcgg 574
Db 301 CCAATGGTGCACCGGATGTGGGCGACATTACTTCATGATGCGCCATGAATCCAGCGG 360
QY 575 aagtggtcaagcagaggtgcatgatacaactcaccataccaccggtgacagactgtg 634
Db 361 AAGTGTCAAGCAGAGGTGAGATGTACAACTCGCGGTACCACCGCTGACAGACTGTG 420
QY 635 tacggcgagtgtygcgaatgaagggcgggggcctttaccgcagctacacacccagc 694
Db 421 TTCGGCGCAGTGTGGCAAAATGAAGGGCGGGCGCTTTTACCGCAGCTACACGACCCAGC 480
QY 695 tgaccatgaacttctcttccaaagccattcacttcacttcagcttcgaacccgagc 754
Db 481 TGACCATGAATGTCCCTTCCAAAGCCATTCACTTTCATGACCTATGAGTTCTCTGCAAGC 540
QY 755 actttaaccccccagagcgg 774
Db 541 ACTTTAAACCCCGAGAGCGG 560

RESULT 15
LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION 60169072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
mRNA sequence.
ACCESSION BE913718
VERSION BE913718.1 GI:104111618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9146 row: o column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo df.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t
```

ORIGIN

```
Query Match 45.9%; Score 502.2; DB 10; Length 598;
Best Local Similarity 92.3%; Pred. No. 3.8e-90;
Matches 551; Conservative 0; Mismatches 43; Indels 3; Gaps 2;

QY 195 gctgccggctgagccactgtcacacgcacatggtggcagcgccgtggcagggatcct 254
Db 1 GCTGCCGGCTGGAGCCACTGTACACACGCACATGTTGGCGGGCGCGTGGCAGGGATCCT 60
QY 255 gaggaactcgtgatgtaccccatcactgcgtcaagaccccgatgcagagtctacagcc 314
Db 61 GGAGCATTTGCGTATGATACCCGATCGCTGCTCAAGACCCCGATGCAGAGCCTACAGCC 120
QY 315 tgacccagctgccgctatcgcaatgtgtgagggccctctgagagtattataagaacga 374
Db 121 TGACCCAGCGCGCGCTATCGGAACGTGTGGAGGCTCTCTGGAGAAATCATGAGGACAGA 180
QY 375 gggcctatgagggcccatgaggggctgaacgtcacagcaacagcgcgaggcctgcccc 434
Db 181 GGGCCTGTGGAGCCCATGCGGGGCTGAACGTTCACAGCAACAGCGCGGGGCTGCCCA 240
QY 435 cgocctttatttgcctgtacgaaagttaaaaaagacattgagtgatgaatccaccc 494
Db 241 CGCCCTCTATTTTGCCTGTCTACGAAAAGTTAAAAAGATTAAGTGACGTAAATCCACCC 300
QY 495 tgggggcaatagccatattgccaatggtggcgccgggtgtgtgcaacattacttcata 554
Db 301 AGGGGCAATAGCCATATTGCCAATGGTCAGCCGGATGTGTGGCGCATTAATCTTCATGA 360
QY 555 tgcagccatgaacctgcggaagtgtcaagcagagatgcagatgtacaactccacata 614
Db 361 TGCAGCCATGAATCCAGCGGAGTCTCTAAGCAGAGAGATGCAGATCTACAACTCGCGTA 420
QY 615 ccacccgggtgacagactgtgtacggcgagtggtggcaaaa--tgaaagggcgccggccttt 672
Db 421 CCACCCGCTGACAGACTGTGTTCGGCAGTGTGTGGCAACACTGACAGGGCGCGGGCTTT 480
QY 673 tacc-gcagctacacacccagctgaccatgaacgttcttccagccattcacttcacat 731
Db 481 TACCTGCAGCTACACGACCCAGCTGACCATGAATGTGCCCTTCCAAAGCCATTCACTTCAT 540
QY 732 gacctatgaattctctcagagacacttttaaccccccagagacggtacacccaagctc 788
Db 541 GACCTATGATTCTCTGCAAGAGCAGCTTTAAACCCCGAGAGCGGTACAAACCCAGTCCC 597
```

Search completed: August 27, 2002, 01:58:09

Job time: 5374 sec
